

# Correcting Storm Displacement Errors in Ensembles Using the Feature Alignment Technique (FAT)

DEREK R. STRATMAN<sup>a</sup>

*NOAA/OAR/National Severe Storms Laboratory, Norman, Oklahoma*

COREY K. POTVIN

*NOAA/OAR/National Severe Storms Laboratory, and Cooperative Institute for Mesoscale Meteorological Studies, and School of Meteorology, University of Oklahoma, Norman, Oklahoma*

LOUIS J. WICKER

*NOAA/OAR/National Severe Storms Laboratory, and School of Meteorology, University of Oklahoma, Norman, Oklahoma*

(Manuscript received 28 November 2017, in final form 27 April 2018)

## ABSTRACT

A goal of Warn-on-Forecast (WoF) is to develop forecasting systems that produce accurate analyses and forecasts of severe weather to be utilized in operational warning settings. Recent WoF-related studies have indicated the need to alleviate storm displacement errors in both analyses and forecasts. A potential solution to reduce these errors is the feature alignment technique (FAT), which mitigates displacement errors between observations and model fields while satisfying constraints. This study merges the FAT with a local ensemble transform Kalman filter (LETKF) and uses observing system simulation experiments (OSSEs) to vet the FAT as a potential alleviator of forecast errors arising from storm displacement errors. An idealized truth run of a supercell on a 250-m grid is used to generate pseudoradar observations, which are assimilated onto a 2-km grid using a 50-member ensemble to produce analyses and forecasts of the supercell. The FAT uses composite reflectivity to generate a 2D field of displacement vectors that is used to align the model variables with the observations prior to each analysis cycle. The FAT is tested by displacing the initial model background fields from the observations or modifying the environmental wind profile to create a storm motion bias in the forecast cycles. The FAT–LETKF performance is evaluated and compared to that of the LETKF alone. The FAT substantially reduces errors in storm intensity, location, and structure during data assimilation and subsequent forecasts. These supercell OSSEs provide the foundation for future experiments with real data and more complex events.

## 1. Introduction

Since Stensrud et al. (2009) first proposed the Warn-on-Forecast (WoF) paradigm, numerous strides have been made in the advancement of storm-scale data assimilation (DA) and numerical weather prediction (NWP) systems. WoF research focuses on using storm-scale (i.e., <4-km

grid spacing), short-term (e.g., 0–3 h) ensemble forecasts to predict severe weather. Stensrud et al. (2013) provide an updated overview of WoF-related experiments and recommend research should focus on challenges associated with data assimilation techniques, ensemble analyses, and model physics parameterizations. Driven by these recommendations, additional WoF-related studies have been completed with focuses on assimilating satellite and radar data (e.g., Jones et al. 2015, 2016; Wheatley et al. 2015; Supinie et al. 2017), assessing the sensitivity of simulated supercells to horizontal grid spacing (e.g., Potvin and Flora 2015), evaluating the sensitivity of simulated supercells to the choice of microphysics parameterization scheme (e.g., Yussouf et al. 2013; Wheatley et al. 2014; Dawson et al.

---

<sup>a</sup> Current affiliation: Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, and NOAA/OAR/National Severe Storms Laboratory, Norman, Oklahoma.

*Corresponding author:* Dr. Derek R. Stratman, derek.stratman@noaa.gov

2015; Stratman and Brewster 2017), and designing and testing a prototype ensemble-based WoF analysis and forecast system (e.g., Wheatley et al. 2015; Yussouf et al. 2015; Jones et al. 2016). As more experiments are completed, some common problematic themes have emerged that require further exploration.

One common problem some recent WoF-related studies have noted is that simulated supercells tend to exhibit substantial displacement errors to the north and east (e.g., Xue et al. 2014; Dawson et al. 2015; Yussouf et al. 2015; Supinie et al. 2017). This bias may arise from errors in the model (e.g., grid resolution and physical parameterizations), observations, and data assimilation systems. Displacement errors introduced during data assimilation can produce large errors in storm location and structure in subsequent forecasts, thereby inhibiting issuance of timely and accurate warnings for all facets of severe weather (i.e., tornadoes, large hail, damaging winds, and flooding). Therefore, the mitigation of displacement errors to improve short-term storm forecasts directly supports the WoF project.

One promising technique to alleviate displacement errors during data assimilation is the feature alignment technique (FAT; Nehr Korn et al. 2015), which is similar to the feature calibration and alignment technique (FCA; Hoffman et al. 1995; Hoffman and Grassotti 1996; Grassotti et al. 1999; Nehr Korn et al. 2014), except without an amplitude-correction component. Different variants of the FAT and FCA methods have been successfully used in previous convection forecasting studies. Brewster (2003a,b) effectively reduced storm location errors in analyses and thereby improved convective forecasts out to 3 h using displacement vectors, computed by minimizing a squared-error difference between forecast and observation fields. Hsiao et al. (2010) utilized a tropical cyclone (TC) vortex relocation technique to improve the TC initializations and forecasts at all forecast times, especially during the earlier parts of the forecast period. This TC vortex relocation technique consists of separating the TC circulation from the background flow, relocating the TC circulation to match the observations, and applying the adjusted TC circulation to the background flow.

Nehr Korn et al. (2014) employed the FCA technique from Grassotti et al. (1999) and expanded on the Hsiao et al. (2010) method of adding the 2D displacement field to full model fields without introducing imbalances (which excite undesirable sound and gravity waves) by accounting for varying terrain. Nehr Korn et al. (2014) first tested the FCA technique on forecasts and observations of vertically integrated water vapor for Hurricane Katrina (2005) in an idealized experiment and concluded that the phase corrections substantially

improved the background and forecasts, even though the forecast hurricane vortex still slightly lagged the observations. Additionally, Nehr Korn et al. (2014) used an ensemble of composite reflectivity forecasts to assess the ability of the FCA to characterize the ensemble dispersion by computing sample covariance statistics for the original ensemble and a post-FCA ensemble. They found that the FCA largely corrects for displacement errors and leaves more homogeneous residual errors, which results in the background error covariance being more Gaussian. Instead of using the FCA technique as a standalone, preprocessing step to data assimilation, Nehr Korn et al. (2015) implemented the FAT into the Weather Research and Forecasting data assimilation system (WRFDA; Barker et al. 2012) and aptly named it the displacement WRF data assimilation system (dWRF). Using the Hurricane Katrina (2005) case again, they found that the dWRF alleviated the large displacement errors in the analysis and improved the short-term forecasts.

In all of the aforementioned studies, three-dimensional variational (3DVar) data assimilation was used. The potential benefits of using the FAT with ensemble data assimilation and forecasting methods were not explored until Nehr Korn et al. (2014). A popular and promising ensemble data assimilation technique for storm-scale forecasting is the ensemble Kalman filter (EnKF; Evensen 1994; Snyder and Zhang 2003). Since the FAT and EnKF have yet to be combined, the present study seeks to implement the FAT component of the FCA (from Nehr Korn et al. 2014) into a local ensemble transform Kalman filter (LETKF; Hunt et al. 2007) data assimilation system to explore the potential contributions of the FAT to WoF's goal of improving storm-scale, short-term forecasts of severe convection. In this initial study, we evaluate the FAT by performing observing system simulation experiments (OSSEs) with an isolated supercell using a coupled Cloud Model 1 (CM1; Bryan and Fritsch 2002) and LETKF system. Our version of the FAT is described in section 2, and the OSSE design is detailed in section 3. Results from the OSSEs are shown and discussed in section 4. Finally, a summary and discussion of the results and potential future work is provided in section 5.

## 2. Overview of the FAT

### a. Cost function

Following Nehr Korn et al. (2014), the FAT determines a 2D field of displacement vectors, which is used in this study to adjust the model state variables at all levels, by minimizing the difference between observation and forecast

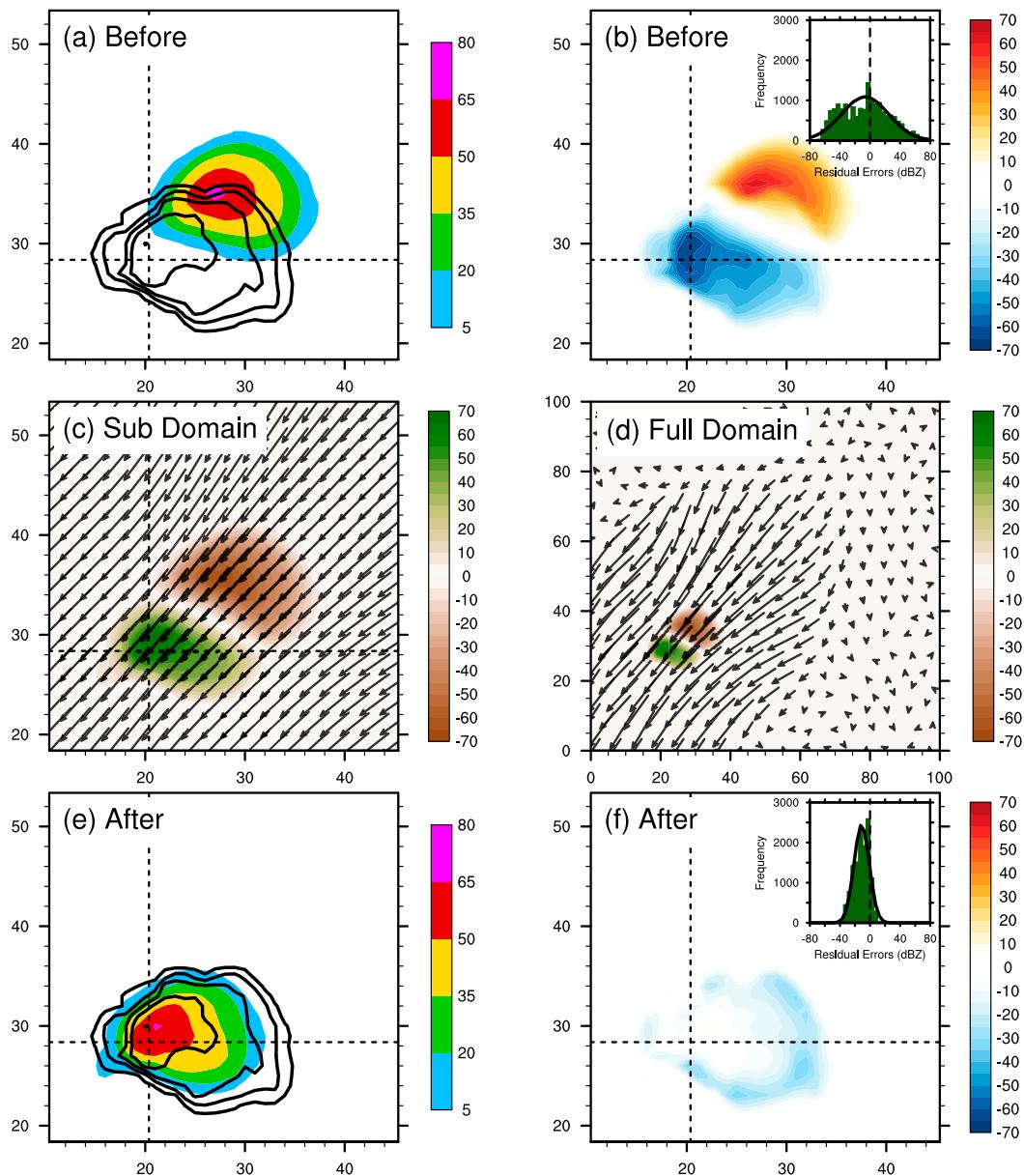


FIG. 1. Example application of the FAT. A 50-member ensemble’s PMM composite reflectivity (dBZ; color shading) and the Truth run’s composite reflectivity (dBZ; black contours) at 5-, 20-, 35-, 50-, and 65-dBZ thresholds are shown for (a) before and (e) after the use of the FAT. Residual errors (color shading) between the ensemble’s PMM composite reflectivity and the Truth run’s composite reflectivity are plotted for (b) before and (f) after the use of the FAT. Inset histograms in (b),(f) depict the number of grid points falling within 5-dBZ error bins (green bars), ideal normal distributions given the mean and standard deviations of the distributions (black solid line), and zero-bias line (black dashed line). Average displacement vectors retrieved from the FAT and the corresponding change in the ensemble’s PMM composite reflectivity (dBZ; color shading) are shown for the (c) subdomain and (d) full domain. The intersection of the horizontal and vertical black dashed lines indicates the location of the Truth run’s maximum composite reflectivity.

fields in model gridpoint space by nonlinearly optimizing the cost function

$$J = J_r[\mathbf{y}, H(\mathbf{x}), \mathbf{a}, \mathbf{b}] + J_p(\mathbf{a}, \mathbf{b}), \tag{1}$$

where  $J_r$  is the residual cost function,  $J_p$  is the penalty cost function,  $\mathbf{y}$  is the set of observations interpolated into gridpoint space,  $H$  is the observation operator,  $\mathbf{x}$  is the model forecast field, and  $\mathbf{a}$  and  $\mathbf{b}$  are the displacement

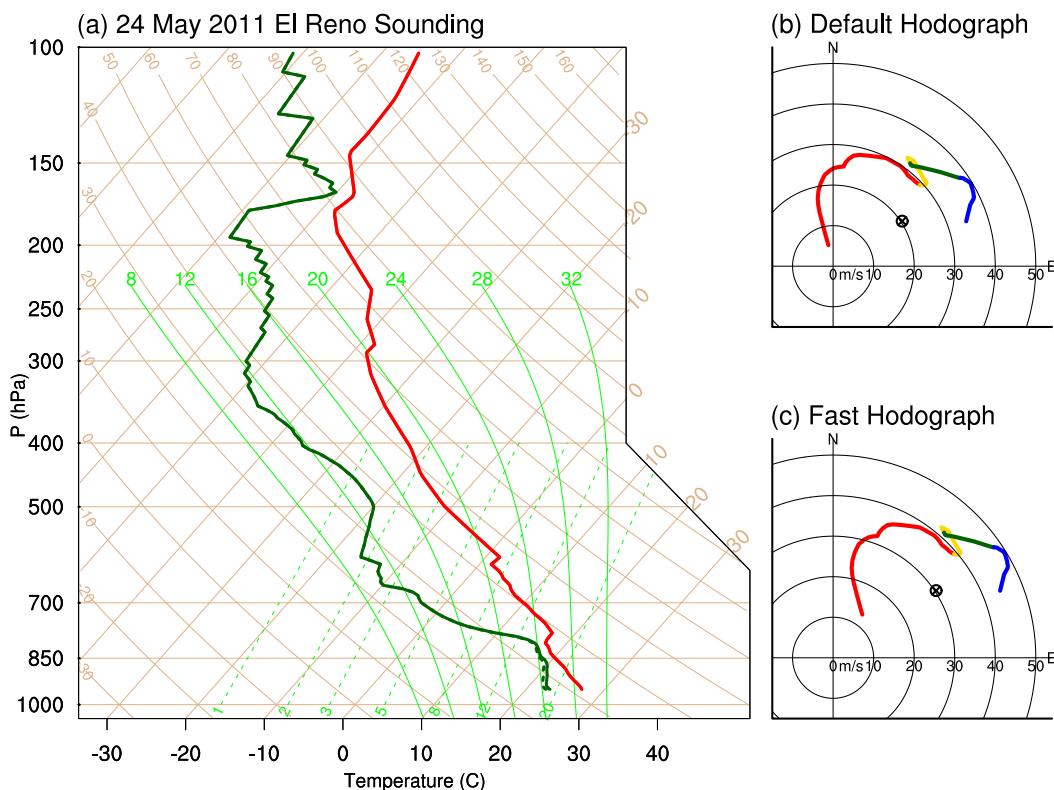


FIG. 2. (a) Skew  $T$ -log $p$  diagram of temperature (red line), default dewpoint temperature (solid green line), and adjusted dewpoint temperature used in some experiments (dashed green line) for the 24 May 2011 El Reno sounding. Hodographs for the (b) default vertical wind profile and (c) fast vertical wind profile used in some experiments are depicted for the 0–3- (red), 3–6- (yellow), 6–9- (green), and 9–12-km (blue) layers. The black markers represent the estimated right-moving storm motion vectors.

vectors in the  $i$  (east–west grid indices) and  $j$  (north–south grid indices) directions, respectively. [Note that  $\mathbf{a}$  and  $\mathbf{b}$  are labeled  $\delta\mathbf{i}$  and  $\delta\mathbf{j}$ , respectively, in [Nehrkorn et al. \(2014\)](#).] For this study, the observation field is composite reflectivity, so  $H$  consists of first computing the 3D reflectivity field using microphysics-scheme-dependent reflectivity equations (see the CM1 source code). Next, the column-maximum reflectivity at every horizontal grid point is determined to yield the forecast composite reflectivity field. The residual cost function is the sum of squared differences between the observation and forecast fields across  $N_o$  observations in gridpoint space weighted by the observational error variance  $\sigma_o^2$  and is formally written as

$$J_r = \sum_{n=1}^{N_o} \frac{[y_n - x_n(i_n + a_n; j_n + b_n)]^2}{\sigma_{o,n}^2}. \quad (2)$$

The penalty cost function constrains the displacements of the forecast field obtained by minimizing the residual cost function alone and is written as

$$J_p = \lambda_s J_s + \lambda_d J_d + \lambda_m J_m + \lambda_b J_b, \quad (3)$$

where  $J_s$  is the smoothness function,  $J_d$  is the divergence function,  $J_m$  is the magnitude function,  $J_b$  is the barrier function, and  $\lambda_s$ ,  $\lambda_d$ ,  $\lambda_m$ , and  $\lambda_b$  are the corresponding weighting coefficients. In [Nehrkorn et al. \(2014\)](#), the smoothness function was not nondimensional like the other cost functions, so for consistency, it is modified to be

$$J_s = \frac{1}{\left[\frac{\Delta d}{(\Delta L)^2}\right]^2} \sum_{n=1}^{N_o} \left( \frac{\partial^2 a_n}{\partial i_n^2} + \frac{\partial^2 a_n}{\partial j_n^2} \right)^2 + \frac{1}{\left[\frac{\Delta d}{(\Delta L)^2}\right]^2} \sum_{n=1}^{N_o} \left( \frac{\partial^2 b_n}{\partial i_n^2} + \frac{\partial^2 b_n}{\partial j_n^2} \right)^2, \quad (4)$$

where  $\Delta d/(\Delta L)^2$  is a representative scale difference between displacements  $\Delta d$  over scale length  $\Delta L$  squared. The smoothness function limits the rate at

which the displacements can vary in space. The divergence constraint penalizes extreme cases of divergence and convergence of displacement vectors to prevent unrealistic adjustments to model field gradients and is written as

$$J_d = \sum_{n=1}^{N_o} \left( \frac{\partial a_n}{\partial i_n} + \frac{\partial b_n}{\partial j_n} \right)^2. \quad (5)$$

The magnitude constraint suppresses displacements in data-void areas to avoid unnecessary adjustments of model fields outside of areas with observations and is written as

$$J_m = \sum_{n=1}^{N_o} \left( \frac{a_n}{S} \right)^2 + \sum_{n=1}^{N_o} \left( \frac{b_n}{S} \right)^2, \quad (6)$$

where  $S$  is a representative displacement scale that penalizes displacement vectors larger than itself. Finally, the barrier constraint severely penalizes displacements larger than  $S$  to prevent excessively large adjustments to model fields and is written as

$$J_b = \sum_{n=1}^{N_o} \left( \frac{a_n}{S} \right)^{20} + \sum_{n=1}^{N_o} \left( \frac{b_n}{S} \right)^{20}. \quad (7)$$

Except for the divergence and barrier penalty functions, [Nehrkorn et al. \(2014\)](#) and other studies using FAT/FCA (e.g., [Hoffman and Grassotti 1996](#); [Grassotti et al. 1999](#)) performed the minimizations in spectral space. For this study, we minimize the residual and penalty cost functions in gridpoint space using finite differencing to calculate their gradients.

*b. Application to composite reflectivity fields*

For this study, simulated observed and forecast composite reflectivity are used in determining the 2D field of displacement vectors from the FAT. The simulated observed composite reflectivity is upscaled from a higher-resolution grid (i.e.,  $\Delta x = 250$  m) to the coarser forecast grid (i.e.,  $\Delta x = 2$  km in our study) using Cressman interpolation. Because the minimization in our version of the FAT is not being done in spectral space, which would allow for truncating higher-wavenumber modes, as in previous studies (e.g., [Hoffman and Grassotti 1996](#); [Grassotti et al. 1999](#); [Nehrkorn et al. 2014](#)), and since these experiments are being completed on convection-allowing scales (i.e., 2-km grid spacing), two smoothing methods are employed to avoid noisy displacement vector fields and resulting unrealistic morphing of storm structures. First, the number of observations, which exist at every point in the domain grid for our study, is thinned by a factor of 2. This thinning

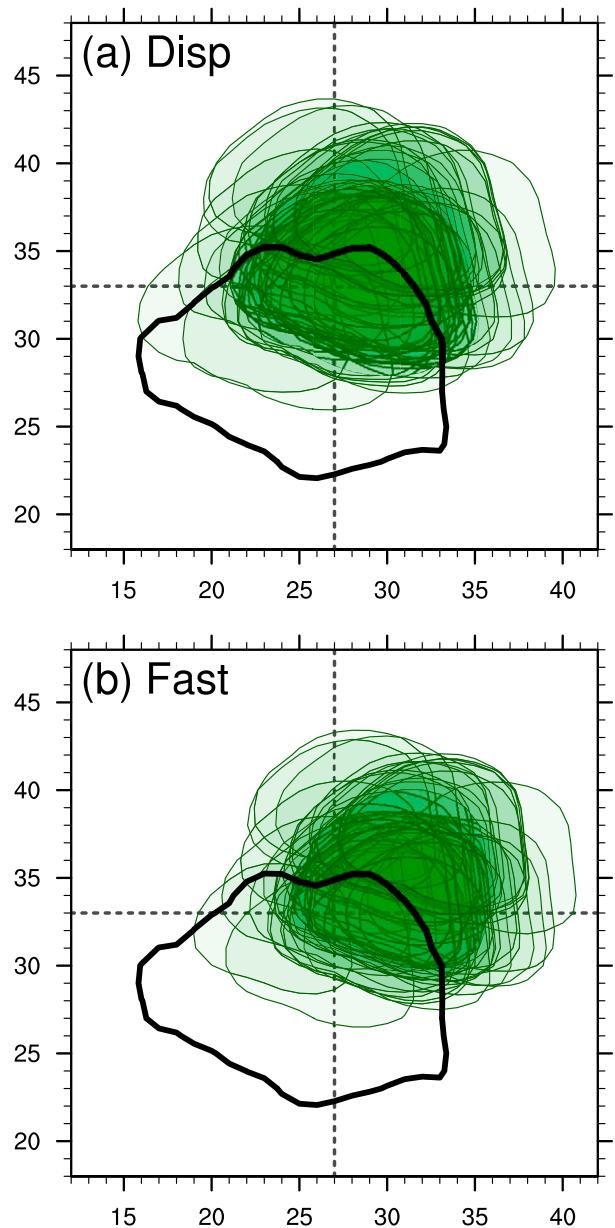


FIG. 3. The 30-dBZ contours of composite reflectivity for the Truth run (black lines) and individual ensemble members (green contours with transparent green shading) at  $t = 30$  min for the (a) Disp and (b) Fast runs. The intersection of the horizontal and vertical black dashed lines provides a reference point and has the same location in both plots.

not only smooths the retrieved displacement vector fields, but also substantially lessens the computation time. Displacement vector fields largely remain similar when using up to every fifth grid point (not shown), so some sacrifice in detail can be made for additional reductions in computation time. Second, a Gaussian filter with  $15\Delta x$  kernel width (e.g., the standard deviation) is applied to the observed and forecast

TABLE 1. Experiment names and their associated descriptions.

Experiment labels	Description
DispTh	Initial displacement of thermal bubbles and Thompson microphysics scheme for 2-km ensembles
DispLG	Initial displacement of thermal bubbles and LFO-Goddard microphysics scheme for 2-km ensembles
FastTh	Faster storm-motion wind profile and Thompson microphysics scheme for 2-km ensembles
DispAdjqvLG	Initial displacement of thermal bubbles, decreased moisture in lowest 2 km, and LFO-Goddard microphysics scheme for 2-km ensembles
Truth	250-m truth simulation used for pseudoradar observations and verification
NoDA	2-km ensemble with no DA
NoFAT	2-km ensemble with only LETKF DA
FAT90	2-km ensemble with LETKF DA and the FAT, which uses the 90% smoothing threshold
FAT50	2-km ensemble with LETKF DA and the FAT, which uses the 50% smoothing threshold
FAT3gp	2-km ensemble with LETKF DA and the FAT, which uses the 90% threshold and is only applied when the initial displacement vector is $\geq 3$ grid points
FAT1st	2-km ensemble with LETKF DA and the FAT, which uses the 90% threshold and is only applied prior to the first DA

composite reflectivity fields. Kernel widths greater than  $10\Delta x$  produce similar results, while smaller kernel widths undesirably morph the storm (not shown). Prior to filtering, the observed and forecast composite reflectivity fields are thresholded such that values below a prescribed percentage (e.g., 90% and 50% are used in this study) of the maximum values are set to zero. We use this thresholding to more closely align the locations of the Gaussian-filtered maximum values and the unfiltered maximum values. Next, the smoothed observed and forecast fields are normalized such that the maximum values in those smoothed fields are the same as the maximum value in the unsmoothed observed field. This normalization helps prevent undesirable storm contraction or expansion.

In the adjoint of  $J_r$  (not shown),  $\mathbf{a}$  and  $\mathbf{b}$  are used to calculate the intervals over which their adjoint equations are calculated:  $\Delta\mathbf{a} = 0.1\mathbf{a}$ ,  $\Delta\mathbf{b} = 0.1\mathbf{b}$ . Having  $\mathbf{a}$  and  $\mathbf{b}$  in the denominator precludes initializing all the elements of  $\mathbf{a}$  and  $\mathbf{b}$  to zero. Instead, they are set to random values between  $-0.25$  and  $0.25$ , except on the

edge of the domain, where they are initialized to zero. To reduce the FAT computation time, a portion of the initialization vectors are replaced with an improved first guess as follows. First, observed and forecast composite reflectivity values  $< 0.1$  dBZ are masked. Then, the centers of mass for both the unsmoothed observation and forecast fields are located. Next, the displacement vector between the two centers of mass is calculated by determining  $\mathbf{a}$  and  $\mathbf{b}$  in the  $i$  and  $j$  directions, respectively. Finally, this displacement vector replaces the original  $\mathbf{a}$  and  $\mathbf{b}$  initializations at grid points where the smoothed observed field exceeds 5 dBZ. Values as large as 60 dBZ can be used for this threshold, but the 5-dBZ threshold generally results in smaller residual errors [e.g., root-mean-square errors (RMSEs) ranging from  $\sim 15$  to  $\sim 19$  dBZ] and nearly constant vectors across the unsmoothed forecast storm (not shown). As a final step in initializing the displacement vectors, any vectors extending outside the domain are shortened to terminate within the domain while preserving the direction of the vectors.

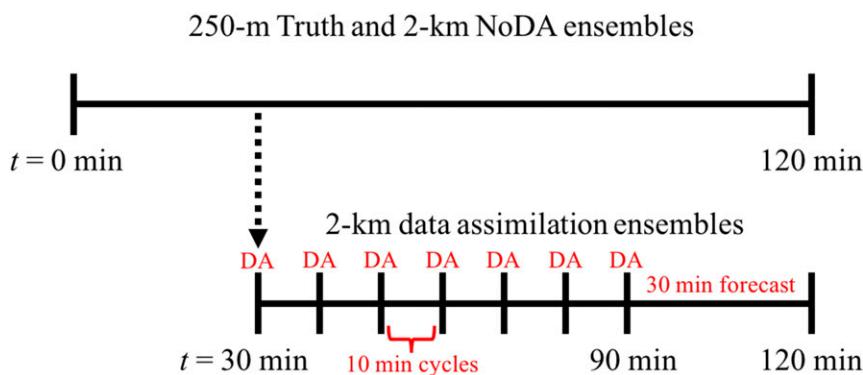


FIG. 4. Timelines for the (top) 2-h Truth and NoDA ensemble simulations and (bottom) 2-km DA experiments, which include six 10-min cycles followed by a 30-min forecast.

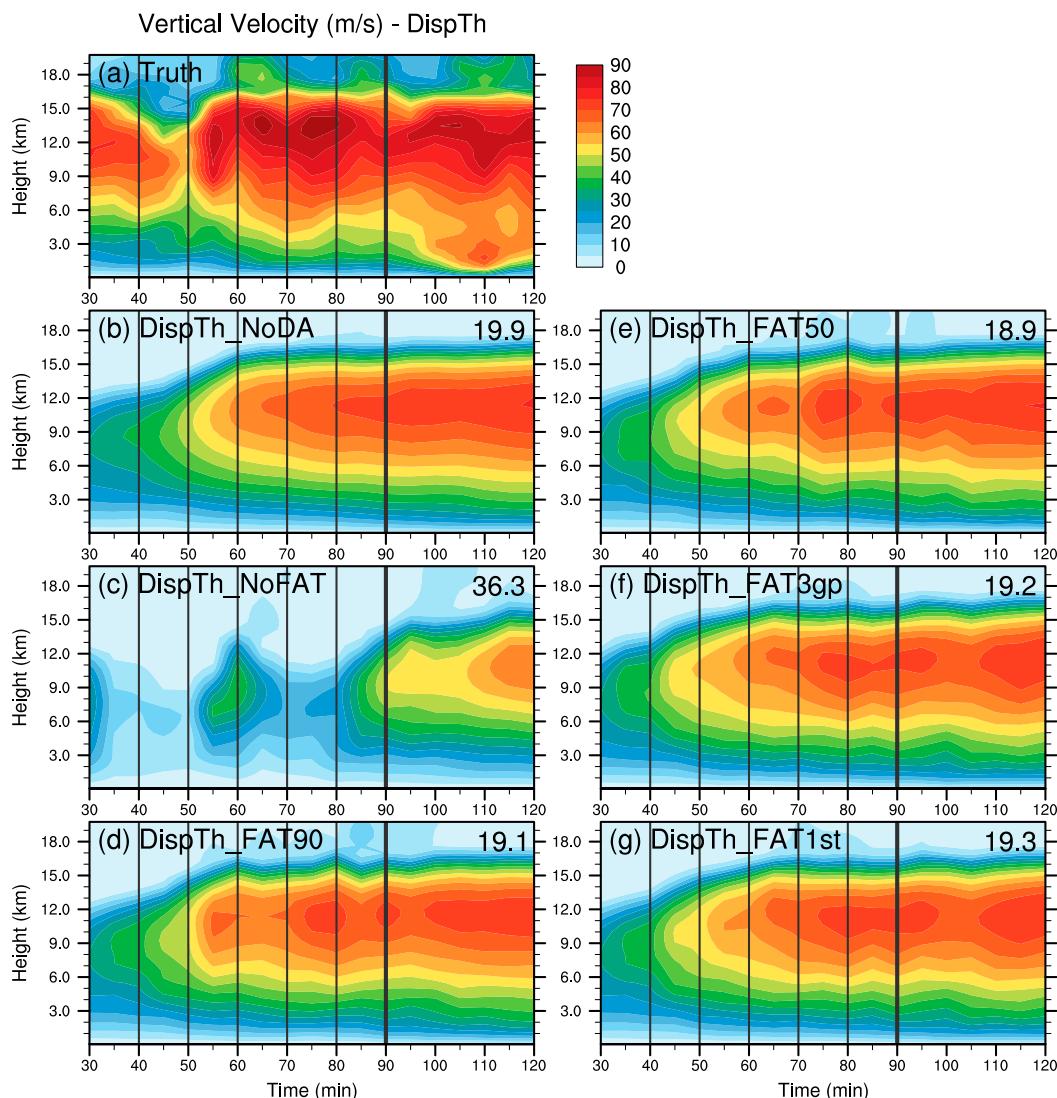


FIG. 5. Time–height plots of ensemble-mean maximum vertical velocity ( $\text{m s}^{-1}$ ) for the DispTh runs. Vertical gray lines represent the times when DA is performed; the thicker gray line indicates the time of the final analysis and initialization of the 30-min forecast. (b)–(g) The RMSEs ( $\text{m s}^{-1}$ ) between the Truth run’s maximum vertical velocities and the DispTh runs’ maximum vertical velocities are annotated in the upper-right corners.

Upon obtaining the smoothed composite reflectivity fields and the first-guess displacement vectors, the optimum **a** and **b** are calculated by minimizing  $J$  [Eq. (1)] using a nonlinear conjugate gradient method (Polak and Ribière 1969). For all our experiments,  $\sigma_o = 7.5$  dBZ,  $S = 50.0$  grid points,  $\Delta d = 1.2$  grid points,  $\Delta L = 1.0$  grid point,  $\lambda_s = 1.0$ ,  $\lambda_d = 0.5$ ,  $\lambda_m = 0.1$ , and  $\lambda_b = 1.0$ . We tested these parameters’ sensitivities by varying their values (i.e.,  $\sigma_o \in [1, 10]$  dBZ,  $S \in [5, 200]$  grid points,  $\Delta d \in [0.1, 10]$  grid points,  $\lambda_s \in [0, 100]$ ,  $\lambda_d \in [0, 1000]$ ,  $\lambda_m \in [0, 100]$ , and  $\lambda_b \in [0, 1000]$ ), while holding the other parameters constant. Through this extensive testing (not

shown), we determined these values are suitable for the isolated storm scenario explored herein. Since the retrieved displacement vectors are on the thinned grid, the displacement vectors are interpolated to the full domain grid before being used to align the forecast fields. As before, vector magnitudes are reduced to prevent vectors from extending outside the domain. It is worth noting that the displacement vectors inward from the domain boundaries are quite similar to the retrieved displacement vectors for a storm centered in a much larger domain (i.e., the domain boundaries have little impact on the displacement vectors near and

around storms). Finally, the 2D field of displacement vectors is applied to all model state variables at all vertical levels.

### c. Example of FAT impact

To depict the operation and impact of the FAT, we illustrate results from an experiment in which a 50-member ensemble forecast of an isolated supercell thunderstorm was displaced  $\sim 14$  km northeastward of the true storm (Fig. 1a). Instead of using the ensemble mean composite reflectivity field, the ensemble probability-matched mean (PMM; Ebert 2001) composite reflectivity field is used to provide an ensemble mean composite reflectivity field with similar values as each individual ensemble member's composite reflectivity field. The initial storm displacement results in a dipole of residual (i.e., truth – each ensemble member) composite reflectivity errors and a non-Gaussian error distribution (Fig. 1b). The FAT determines the optimum 2D field of displacement vectors for each ensemble member forecast and true composite reflectivity fields (Figs. 1c,d). After the model state fields are adjusted by the displacement vectors, the ensemble PMM storm is now collocated with the truth storm (Fig. 1e), greatly reducing the residual errors and causing them to be more Gaussian (cf. Figs. 1b,f), as Nehr Korn et al. (2014) also demonstrated. Most data assimilation methods, including EnKF, assume background forecast and observation errors have a Gaussian-shaped probability density function (PDF), so a more Gaussian distribution of background forecast errors results in a more optimal EnKF analysis (Xue et al. 2007).

## 3. Experiment design

Since supercells produce much of the significant severe weather in the United States, this initial study focuses on the FAT's impact on analyses and forecasts of isolated supercells. An OSSE framework is designed to evaluate the FAT and its sensitivities using both perfect- and imperfect-model experiments. Model simulations are performed using version cm1r18.3 of CM1 (Bryan and Fritsch 2002). In OSSEs with data assimilation, we use an LETKF system developed at NSSL (Thompson et al. 2015) from the Miyoshi (2011) LETKF code. Both the Thompson (Thompson et al. 2008) and LFO-Goddard (Braun and Tao 2000) microphysics schemes are utilized in these experiments. Other model details include vertically stretched grid with  $dz$  around 100 m near the surface and 700 m above 8.2 km, 25-km domain top, 3-s time step, open-radiative lateral boundary conditions, free-slip bottom and top boundary conditions, fifth-order horizontal and vertical advection schemes,

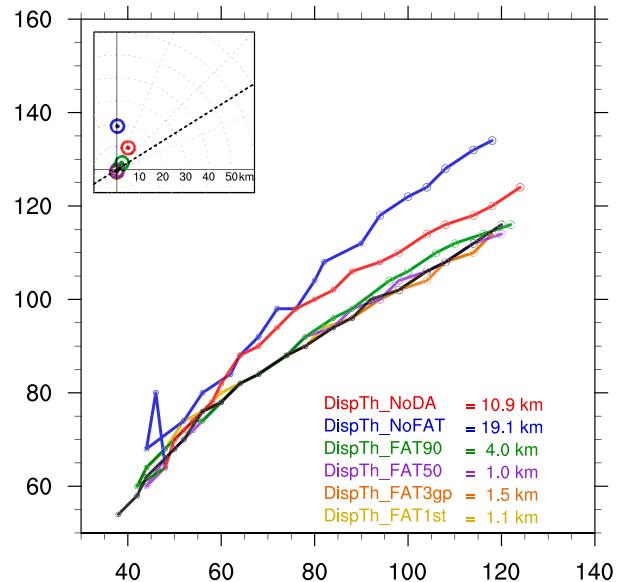


FIG. 6. Maximum-UH tracks (solid lines) and locations (circle markers) are shown every 5 min from 30 to 120 min for the Truth (black) and DispTh ensemble runs. The UH locations in the latter are determined using the ensemble probability-matched means for UH. Small (large) circle markers designate UH locations during the 60-min DA (30-min forecast) period. Average location errors during the forecast period are annotated in the bottom right. (top left) The inset plot depicts the average locations of the forecast UH centers relative to the Truth run's UH centers at the same times. For reference, the Truth run's average storm motion direction during the 30-min forecast period is represented by the black dashed line.

weighted essentially nonoscillatory (WENO; Shu 2003; Shen and Zha 2010) scheme for advecting scalars on the final Runge–Kutta step only, 1.5-order TKE closure for LES (Deardorff 1980), and Rayleigh damping for levels above 20 km.

A 2-h simulation of a supercell (Truth) is computed on an  $800 \times 800 \times 50$  gridpoint domain with 250-m grid spacing. The Truth run's simulated supercell is initiated with a 5-K warm bubble (10 km across and 1.5 km deep) placed at 1.5 km AGL in a homogeneous environment provided by the 24 May 2011 "El Reno" temperature, moisture, and wind vertical profiles (Figs. 2a,b; same as in Tanamachi et al. 2015). The Truth run's thermal bubble is initially located at  $x = 20$  km,  $y = 20$  km (origin located at lower-left corner of domain). Output from the Truth run is used to generate the pseudo-radar observations (reflectivity and radial velocity) at each 250-m grid point. The pseudo-observations are interpolated onto a 2-km,  $100 \times 100 \times 50$ -point analysis grid (that is collocated with the 250-m grid) using a Cressman weighting function with  $\sim 2800$ -m radius of influence. The pseudoradar observations are then output to an observation sequence file using the Data

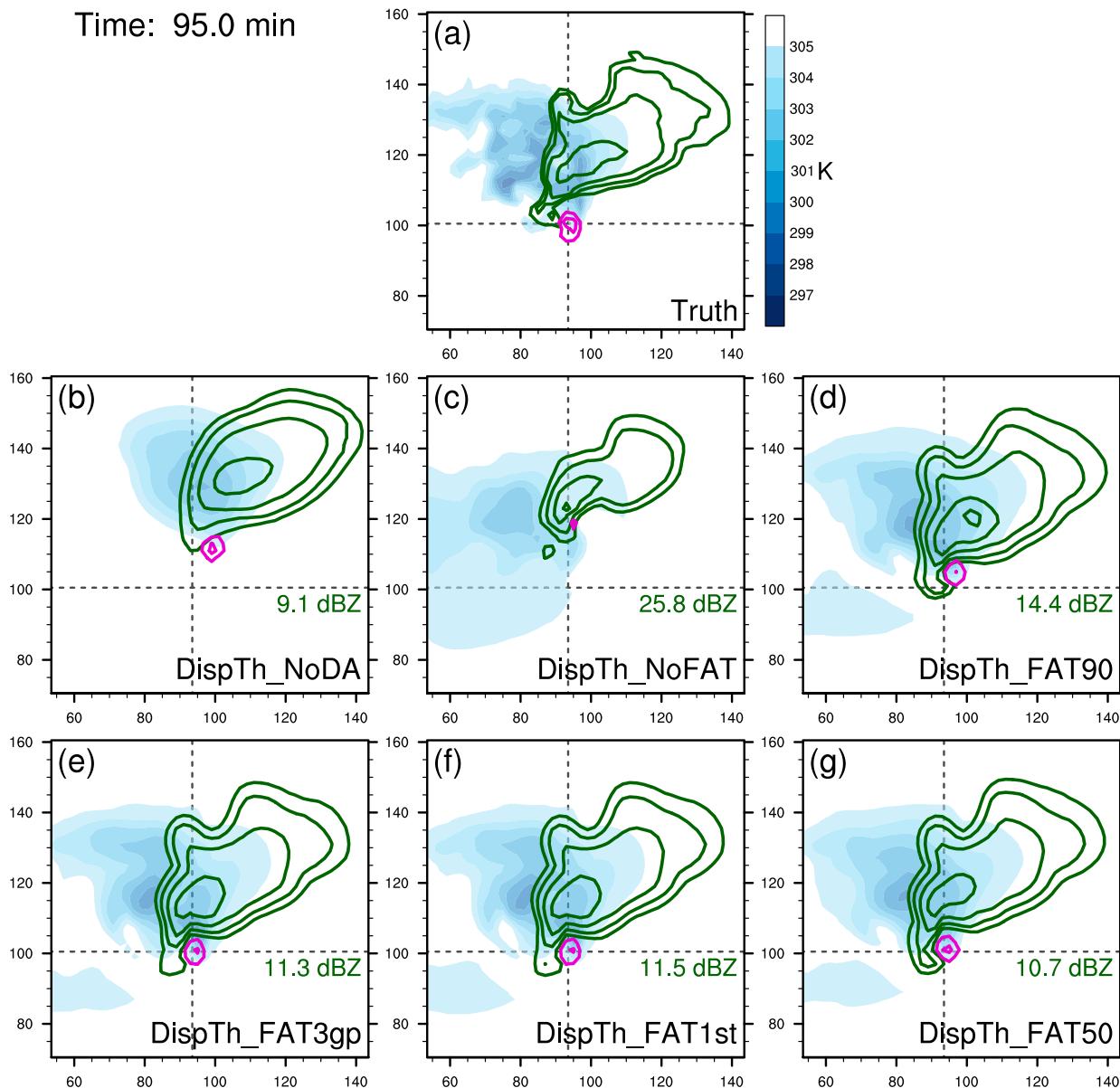


FIG. 7. The 1-km simulated reflectivity at 20, 30, 40, 50, and 60 dBZ (green contours), 2–5-km UH at 500 and 2000  $m^2 s^{-2}$  (magenta contours), and potential temperature (K) at the lowest model level (blue shading) are depicted for the (a) Truth and (b)–(g) various DispTh ensemble runs valid at  $t = 95$  min. Note that the PMM fields are used for the ensembles’ output. The RMSEs for the PMM simulated reflectivity fields are annotated in green in (b)–(g). The intersection of the horizontal and vertical dashed lines represents the Truth run’s max-UH location.

Assimilation Research Testbed (DART; Anderson et al. 2009) format so that they can be assimilated into the CM1–LETKF system. Note that the radial velocity field assumes the radar is located at the center of the domain at 0 m AGL.

Four sets of data assimilation experiments are conducted, all with 50-member ensembles. For each experiment set, an ensemble of simulations is produced to provide the background fields for the first analysis.

These no-data-assimilation (NoDA) simulations initialize convection with a thermal bubble having the same shape as in the Truth run, but with a 6-K perturbation. To create ensemble spread, the bubble locations are randomized by adding a set of perturbations randomly drawn from a Gaussian distribution (standard deviation = 4 km) to the mean bubble location. The Truth run uses the Thompson two-moment microphysics scheme (Thompson et al. 2008). To test the FAT with

Time: 120.0 min

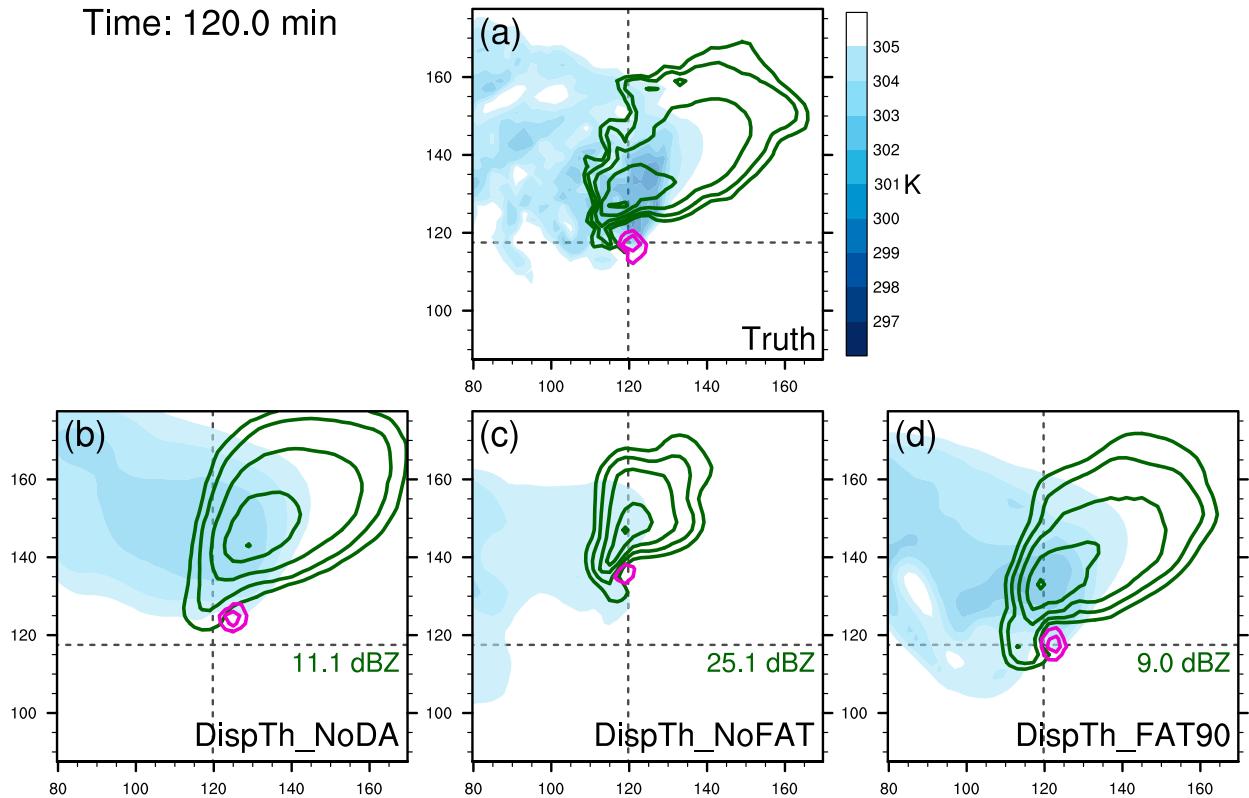


FIG. 8. As in Fig. 7, but only for the NoDA, NoFAT, and FAT90 ensembles at  $t = 120$  min.

both perfect- and imperfect-model physics, two sets of experiments use the Thompson scheme (Th), while the other two use the LFO-Goddard microphysics scheme (LG; Braun and Tao 2000).

Two of the four sets of data assimilation experiments use background fields from 2-km NoDA ensembles, which use a perfect environment and have thermal bubbles inserted  $\sim 14$  km northeastward (10 km eastward and northward) of the Truth run's bubble location (DispTh and DispLG; Fig. 3a; Table 1). Another set of experiments uses background fields from 2-km NoDA ensembles that use a vertical wind profile that produces  $10 \text{ m s}^{-1}$  faster storm motion than the default wind profile (Fig. 2c) and mean thermal bubbles collocated with the Truth run's bubble (FastTh; Fig. 3b; Table 1). Finally, we conduct a set of experiments (i.e., DispAdjvLG) that are similar to the DispLG experiments, except that the boundary layer in the NoDA ensemble is drier than in the Truth run (Fig. 2a; Table 1). The mixing ratio is adjusted following a sine curve with a 3% decrease at the surface (i.e.,  $\sim 0.5^\circ\text{C}$  decrease in dewpoint temperature) and 0% change at 2 km AGL. This adjustment mimics cases where analysis errors arise from errors in observations and/or boundary layer parameterization schemes (e.g., Coniglio 2012; Coniglio

et al. 2013). In addition to providing the background fields for the data assimilation experiments, the NoDA ensembles provide a baseline for assessing data assimilation, grid spacing (i.e., 250 m vs 2 km), and environment error impacts on the modeled storm evolution.

Most of the data assimilation experiment sets described above comprise the following experiments (Table 1): FAT not applied (NoFAT), FAT applied with the 90% or 50% smoothing thresholds (FAT90 and FAT50, respectively), FAT only applied prior to the first data assimilation cycle (FAT1st), and FAT applied only when the centers of mass of the smoothed observation and forecast fields are  $\geq 3$  grid points apart (FAT3gp). The 90% smoothing threshold is used for both the FAT1st and FAT3gp experiments. For the DispAdjvLG experiments, only the NoDA, NoFAT, and FAT90 simulations are produced. For all data assimilation experiments, the first analysis update is performed at  $t = 30$  min using the NoDA ensemble forecast valid at that time as the background state. This first analysis update is followed by six 10-min data assimilation cycles. Additive noise (Dowell and Wicker 2009) is applied to the wind components (i.e.,  $u$ ,  $v$ , and  $w$ ), potential temperature, and dewpoint temperature (which is then translated into the water vapor mixing

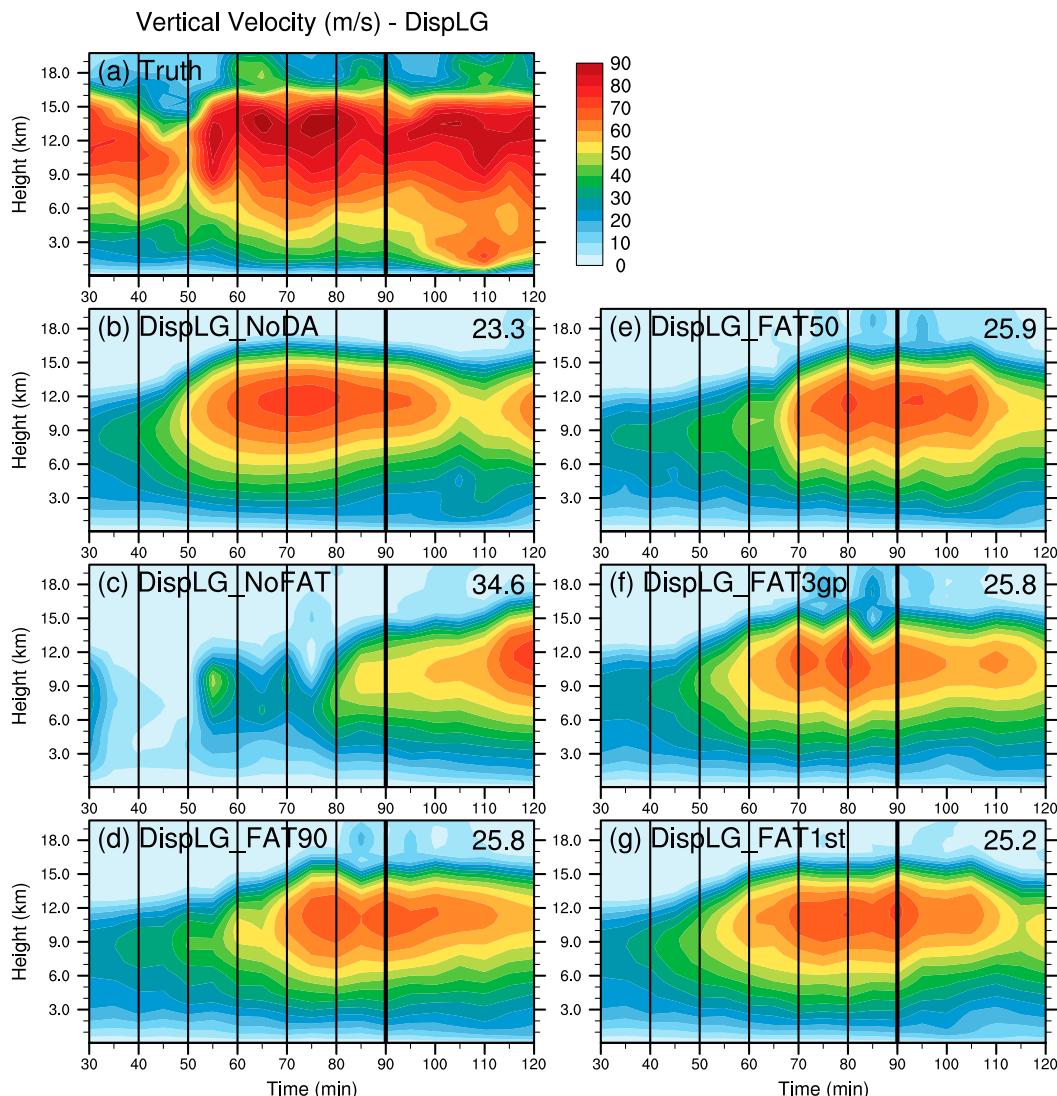


FIG. 9. As in Fig. 5, but for DispLG.

ratio field) fields where reflectivity >30dBZ. In experiments using the FAT, the technique is run immediately before each analysis update. After the final analysis ( $t = 90$  min), forecasts are integrated to  $t = 120$  min (Fig. 4).

**4. Results**

*a. Initial displacement experiments*

1) PERFECT MICROPHYSICS

Unlike the Truth run and DispTh\_NoDA, DispTh\_NoFAT's storm fails to develop a strong updraft [e.g.,  $>50 \text{ m s}^{-1}$ ; note these values are supported by previous observations (e.g., Lehmler et al. 2001)] until after the

final data assimilation cycle (Fig. 5c) and fails to develop a strong low-level updraft (e.g.,  $>25 \text{ m s}^{-1}$  below 3 km) at any time (Figs. 5a–c). In contrast, all of the ensembles that use the FAT produce deep, intense updrafts similar to those in DispTh\_NoDA (cf. Figs. 5b,d–g). Moreover, updraft speeds in these ensembles approach those in the Truth run sooner than those in DispTh\_NoDA do. Also, the depth and intensity of the updrafts in the ensembles using the different FAT variations are quite similar. In addition to these subjective evaluations, RMSEs between the Truth run's maximum vertical velocities and the DispTh runs' average maximum vertical velocities are computed; these agree with the subjective comparisons (cf. Figs. 5b–g). Thus, applying the FAT only to the first

analysis (DispTh\_FAT1st) is sufficient in this experiment to alleviate the errors produced in DispTh\_NoFAT (cf. Figs. 5c,g).

The use of only LETKF in DispTh\_NoFAT yields forecasts with substantially larger average displacement errors during the forecast period than DispTh\_NoDA (i.e., 19.1 vs 10.9 km; Fig. 6). This result stems from the LETKF's inability to properly handle the large, non-Gaussian errors associated with the initial displacement, which contributes to the storm being weaker and remaining elevated above the boundary layer longer. Hence, DispTh\_NoFAT's mean storm location is well to the north of the Truth and DispTh\_NoDA runs' storm. Conversely, the inclusion of the FAT greatly reduces the displacement errors in the other ensembles, with average forecast errors  $\leq 4$  km (Fig. 6).

To highlight the impacts of the different data assimilation methods on storm structure, we now examine ensemble output valid 5 min into the 30-min forecast initialized from the final analysis (i.e.,  $t = 95$  min). At this time, the differing effects of the various data assimilation procedures are still mostly present, while the model effects are beginning to become more apparent. Comparing each ensemble's PMM  $\sim 1$ -km reflectivity (i.e., model level 8), the most noticeable difference is that DispTh\_NoFAT's storm is by far the smallest (Fig. 7c). This result is not attributable to using the PMM technique in forming the ensemble mean fields since there is little ensemble spread, and individual members' storms are similar in size (not shown). For DispTh\_NoDA, the PMM technique causes the individual ensemble members' reflectivity hook echo to become nonexistent for reflectivities greater than 20 dBZ, owing to larger ensemble spread (Fig. 7b). However, the general shape and size of DispTh\_NoDA's PMM storm is similar to the individual members' storms.

In addition to having a smaller storm, DispTh\_NoFAT has a weaker, smaller mesocyclone than the Truth run and the other ensembles per the 2–5-km updraft helicity (UH; Kain et al. 2008) field (Fig. 7c), and its cold pool intensity and structure least resemble those of the Truth run. In contrast, the ensembles incorporating the FAT produce reflectivity, UH, and near-surface potential temperature fields substantially more similar to the Truth run's fields (Figs. 7a,d–g). This result is objectively supported by the RMSEs between the Truth run's 1-km reflectivity field and the ensembles' location-corrected (i.e., using the locations of maximum UH) PMM 1-km reflectivity fields where observed or forecast reflectivity exceed 20 dBZ (Figs. 7b–g). At  $t = 120$  min (i.e., 30-min forecast), the impact from using the FAT is still clearly apparent in the reflectivity and surface cold pool features (cf. Figs. 7b–d, 8b–d). Generally

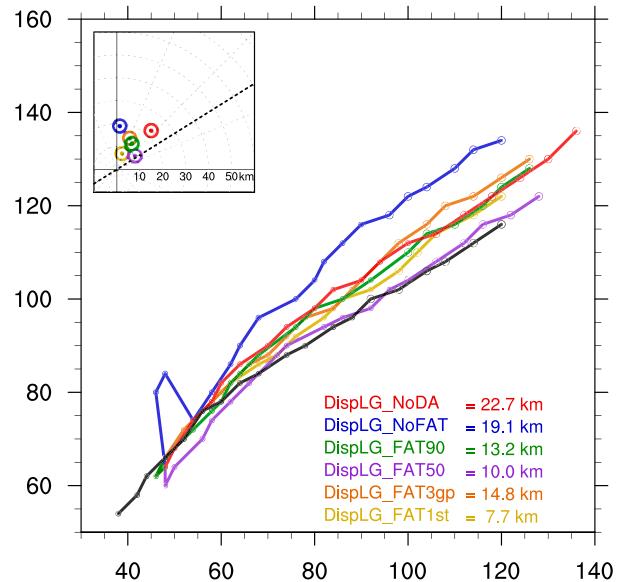


FIG. 10. As in Fig. 6, but for DispLG.

speaking, the FAT in these perfect-model (apart from grid spacing) experiments substantially improves the analyses and forecasts of the supercell.

## 2) IMPERFECT MICROPHYSICS

Unlike DispTh\_NoDA's storm in the perfect-microphysics experiments, DispLG\_NoDA's storm erroneously weakens during  $t = 80$ – $100$  min before re-strengthening during the final 15 min of the 30-min forecast (Fig. 9b). Ideally, the use of data assimilation would mitigate the effects of the microphysical errors and thereby improve the analyzed storm intensity. However, the DispLG\_NoFAT storm struggles to develop an intense, deep updraft during and immediately after the data assimilation window (Fig. 9c). Applying the FAT prior to the analysis updates, on the other hand, generally improves upon DispLG\_NoFAT through  $t = 100$ – $110$  min, though the depth of the intense portion of the updraft erroneously decreases toward the end of the forecast period (Figs. 9d–g). Of all the ensembles, DispLG\_FAT90's updraft best matches the Truth run's updraft evolution by sustaining an intense, deep updraft (e.g., vertical velocities  $> 55 \text{ m s}^{-1}$ ) through the entirety of the forecast period (i.e.,  $t = 90$ – $120$  min; cf. Figs. 9a,d).

While the FAT does not improve storm location as much in these imperfect-microphysics runs as in the perfect-microphysics (DispTh) runs, the location errors are still reduced up to  $\sim 66\%$  when compared to DispLG\_NoDA (Fig. 10). Once again, the LETKF-only experiment, DispLG\_NoFAT, yields a storm that is much farther north than in the other experiments. The DispLG\_NoFAT storm closely tracks the Truth run's

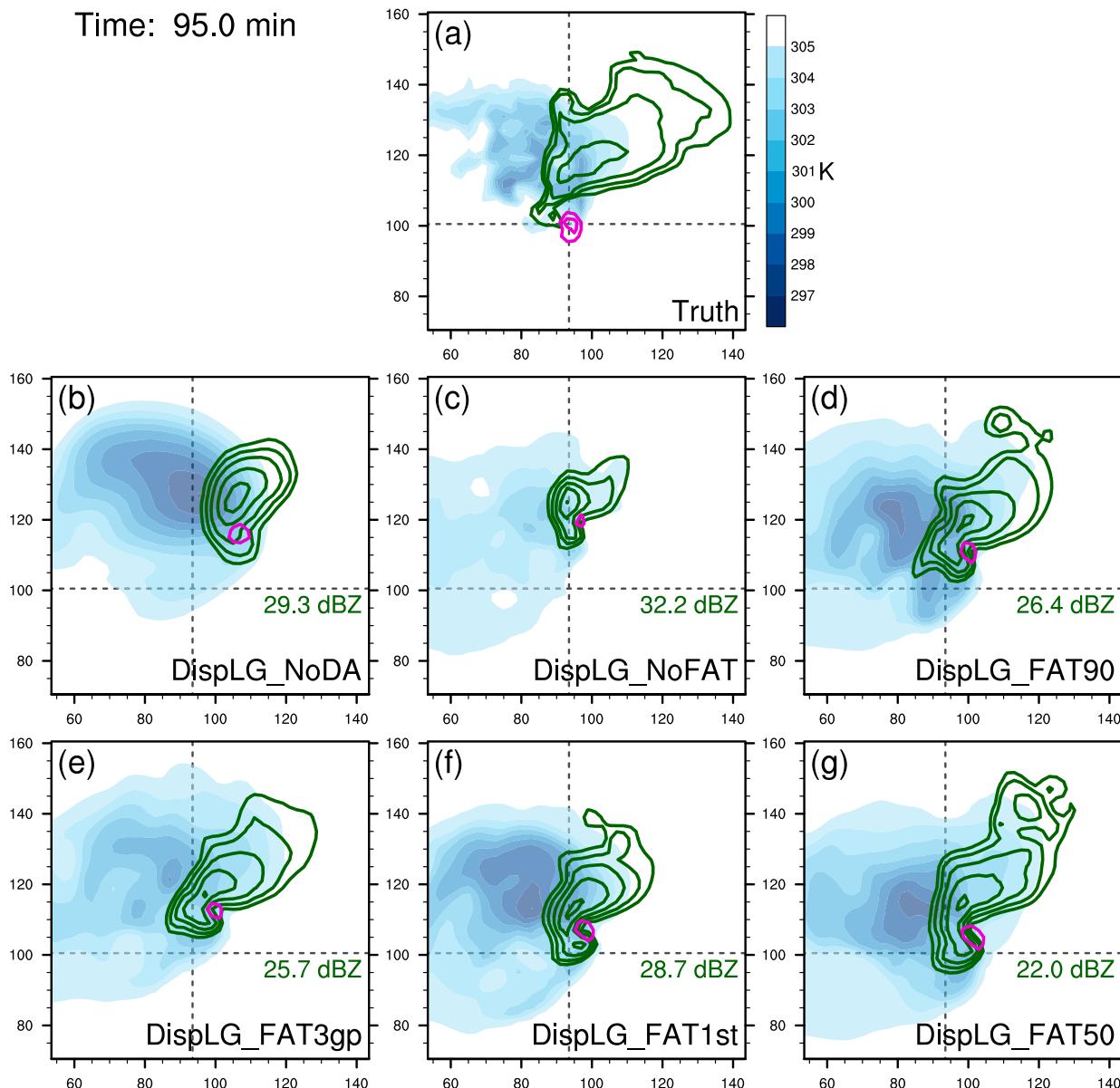


FIG. 11. As in Fig. 7, but for DispNet.

storm in longitude, but as depicted in Fig. 11c, this is likely attributable to the DispNet\_NoFAT storm being weaker, producing a warmer cold pool, and therefore experiencing slower forward propagation, which fortuitously compensates for its initial eastward displacement. As shown in Fig. 10, running the FAT beyond the first analysis cycle can be counterproductive (e.g., DispNet\_FAT90 vs DispNet\_FAT1st). Specifically, the 90% smoothing threshold (DispNet\_FAT90 and DispNet\_FAT3gp) can cause the smoothed reflectivity fields to vary more in location due to the relatively large variability of higher-reflectivity characteristics (e.g.,

location, size, and shape), as compared to using the 50% smoothing threshold (DispNet\_FAT50). This sensitivity can result in undesirable adjustments by the FAT. For example, DispNet\_FAT90 and DispNet\_FAT3gp have larger average forecast errors than the other FAT ensembles mainly due to the final FAT computation increasing the storm displacement error (Fig. 10). Even with this error increase, however, the inclusion of the FAT still reduces storm location errors more than data assimilation alone (DispNet\_NoFAT).

Similar to the DispNet experiments, the storm produced by DispNet\_NoFAT is much smaller than the

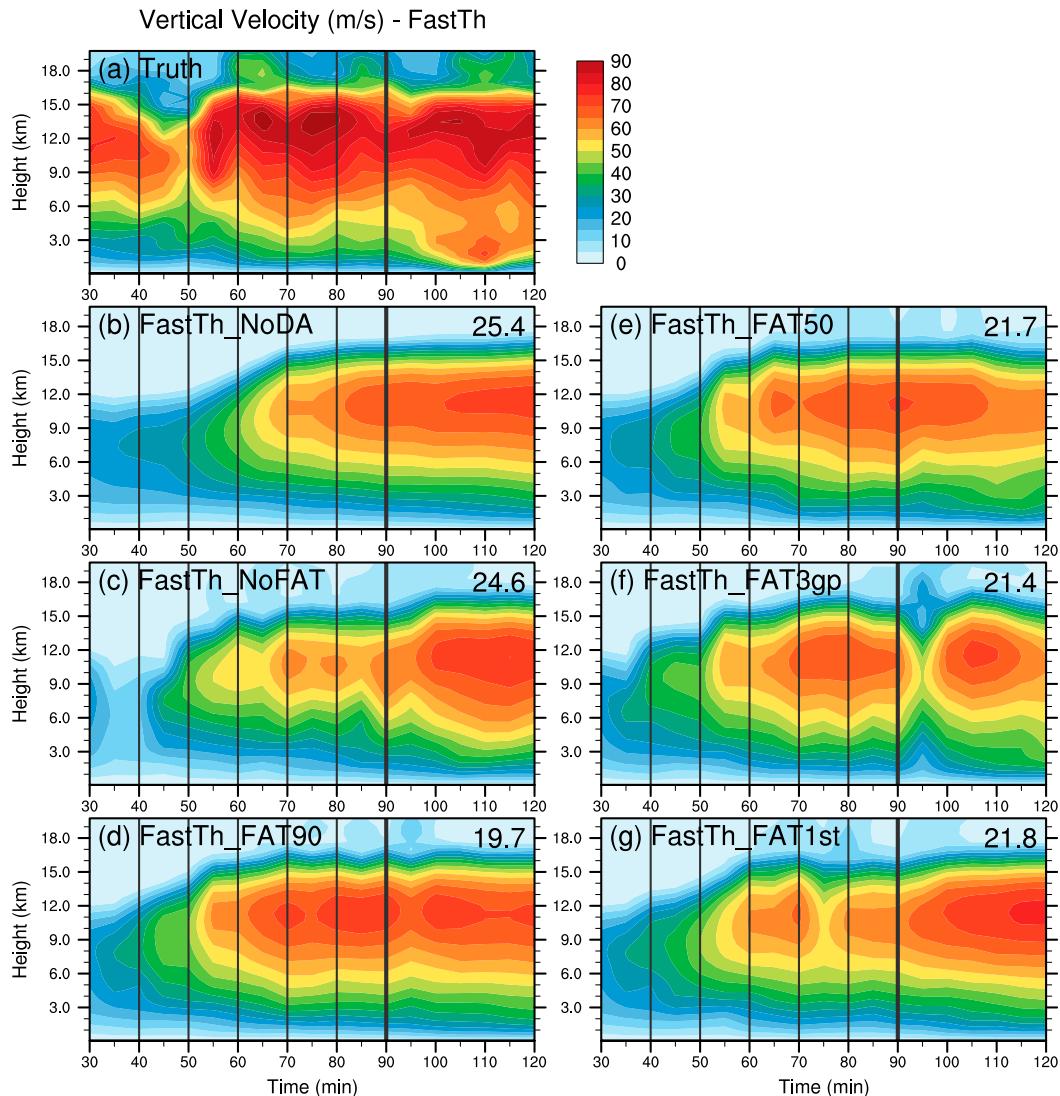


FIG. 12. As in Fig. 5, but for FastTh.

Truth run's storm and the other ensembles' storms, while the ensembles with FAT produce larger storms that better resemble the Truth storm (Fig. 11). The ensembles using the FAT tend to produce weak convection to the north of the supercell, but this convection has negligible impact on the supercell. Since weak convection develops in the Truth run as well (not shown), we do not regard this additional convection as a major fault of the FAT. As in the perfect-microphysics experiments, the updraft in the LETKF-only ensemble is smaller and weaker than in the other ensembles (cf. Figs. 11b–g). The ensembles using the FAT produce cold pools that tend to be a little colder than the Truth's cold pool, similar to DisPLG\_NoDA's cold pool (cf. Figs. 11a,b,d–g). This bias presumably arises largely from microphysics errors and perhaps

also from the coarse grid spacing. As previously mentioned, DisPLG\_NoFAT's cold pool is too warm owing to the storm being weaker (Fig. 11c). Overall, as in the perfect-microphysics experiments, the FAT again produces substantial improvements in storm intensity, location, and structure.

#### b. Fast sounding with perfect microphysics

With the faster vertical wind profile, the ensembles' storms are displaced  $\sim 18$  km northeastward of the Truth run's storm at the time of the first data assimilation (i.e.,  $\sim 4$  km farther away than in the DisPLG experiments; cf. Figs. 3a,b). Since there is now a continuous forcing for storm motion bias, the displacement errors that must be corrected at the end of each forecast cycle are larger than in the perfect-environment experiments. These

experiments, therefore, test the impact of the FAT in the common scenario where a storm motion bias persists through the data assimilation window and into the free forecast. The storm in FastTh\_NoDA (Fig. 12b) takes ~10–15 min longer to develop than the storm in DispTh\_NoDA (Fig. 5b) likely owing to the increased diffusion in the advection terms resulting from the faster wind speeds (Kniewicz et al. 2007). In all of the data assimilation experiments, a strong storm updraft develops ~10 min earlier than in FastTh\_NoDA (cf. Figs. 12b–g). The updraft in FastTh\_NoFAT, however, is erroneously weaker and shallower than in FastTh\_NoDA (cf. Figs. 12b,c). Conversely, FastTh\_FAT90 produces the deepest and generally strongest updraft (Fig. 12d), and the updraft evolution in the remaining ensembles likewise generally better matches the Truth run than does FastTh\_NoFAT, which is evident in the RMSEs as well (Fig. 12). Note that the decrease in FastTh\_FAT3gp’s maximum vertical velocities after the last data assimilation (Fig. 12f) likely originates from the LETKF and not the FAT. This inference was determined by examining the maximum vertical velocity field before and after the FAT and before and after the LETKF data assimilation. Even so, the FastTh\_FAT3gp storm quickly recovers during the forecast.

Unlike in the Disp experiments, FastTh\_NoFAT’s storm track mostly remains south of FastTh\_NoDA’s storm track (Fig. 13), consistent with FastTh\_NoFAT’s storm becoming organized sooner than in DispTh\_NoFAT and DispLG\_NoFAT (cf. Figs. 5c, 9c, 12c). However, FastTh\_NoDA’s forecast-storm track lies north of the Truth run’s storm track and the FAT ensembles’ storm tracks. Displacement errors in the ensembles with FAT applied are reduced up to ~76% over FastTh\_NoDA’s displacement errors. Even running the FAT only once (i.e., FastTh\_FAT1st) reduces the displacement errors ~44%, as compared to only ~15% for FastTh\_NoFAT. Furthermore, all of the ensembles using the FAT yield storm tracks that are more in line with the Truth run’s storm track, except with the fast storm motion bias (Fig. 13).

All of the ensembles with data assimilation, especially FastTh\_FAT90, FastTh\_FAT3gp, and FastTh\_FAT50, have storms with 1-km reflectivity structures closer in shape to the Truth run’s storm than FastTh\_NoDA’s storm through the end of the forecast period (e.g., Fig. 14). Unlike in the Disp experiments, the storm in FastTh\_NoFAT is not substantially smaller than in the Truth run and the other ensembles through the entire forecast period (Fig. 14). However, as in the Disp experiments, FastTh\_NoFAT’s storm is generally less organized than in the other ensembles through most of the data assimilation period (e.g., Fig. 12), presumably

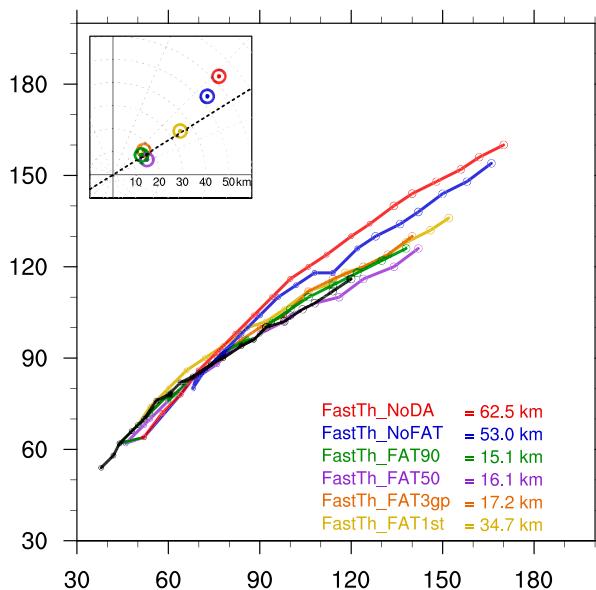


FIG. 13. As in Fig. 6, but for FastTh.

attributable to the negative impact from the LETKF’s handling of the constant introduction of displacement errors (not shown). At  $t = 95$  min, FastTh\_FAT90, FastTh\_FAT3gp, and FastTh\_FAT50 have well-established cold pools whose structure and, less so, intensity resemble that of the Truth’s cold pool (cf. Figs. 14a,d,e,g). FastTh\_NoFAT and FastTh\_FAT1st’s cold pools are weaker but improve through the rest of the forecast period (not shown). Overall, the FAT once again improves analyses and forecasts, this time in the scenario where the model storm continually outpaces the observed storm during the forecast cycles.

*c. Adjusted BL moisture*

To evaluate the impact of the FAT in the presence of another common source of forecast error, some of the imperfect-microphysics experiments were repeated with the moisture in the lowest 2 km AGL of the initialization sounding reduced, as described in section 3. The drier boundary layer causes DispAdjqvLG\_NoDA to fail at sustaining updraft speeds  $>25 \text{ m s}^{-1}$  through most of the 2-h simulation (Fig. 15b). Incorporating no-FAT data assimilation (i.e., DispAdjqvLG\_NoFAT) actually further degrades the storm, except near the end of the forecast period (Fig. 15c). Conversely, the inclusion of the FAT (i.e., DispAdjqvLG\_FAT90) yields maximum vertical velocities greater than  $50 \text{ m s}^{-1}$  during the entire forecast period (Fig. 15d).

As implied by the updraft speed composites, DispAdjqvLG\_NoDA has a small, elevated storm with little rotation at 5 min into the forecast period (Fig. 16b). Even worse, DispAdjqvLG\_NoFAT at this time has two

Time: 95.0 min

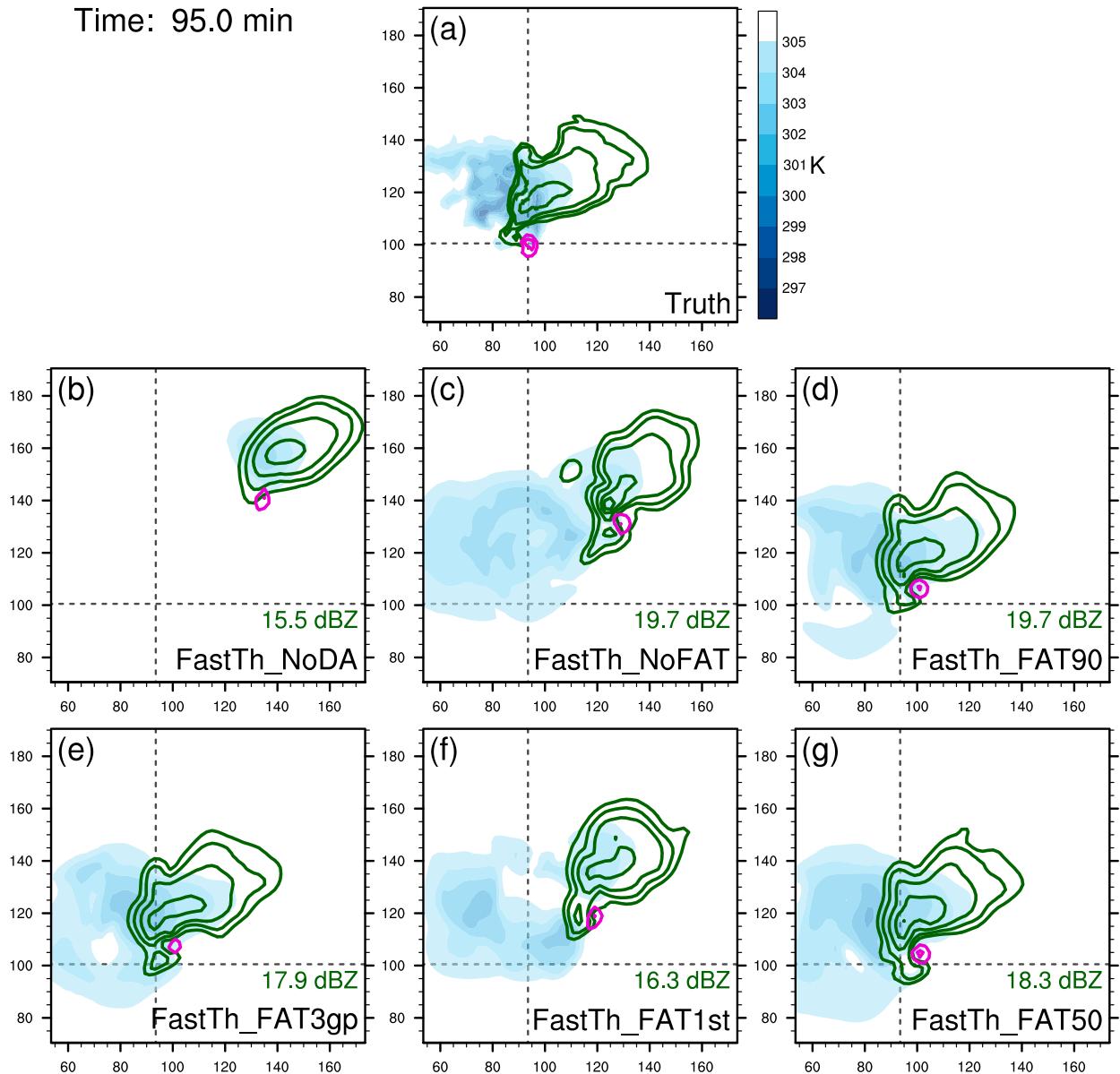


FIG. 14. As in Fig. 7, but for FastTh.

very weak, unorganized convective cells with a cold pool that (not surprisingly) is too warm (Fig. 16c). Dis-pAdjvLG\_FAT90, on the other hand, has a rotating storm with an established cold pool (Fig. 16d). These results indicate that data assimilation by itself can be very sensitive to the combination of storm displacement and environment errors, even to the point of failing to establish a supercell after several analysis cycles. More specifically, the storm displacement errors contribute to non-Gaussianity of errors and large analysis increments, both of which degrade the analysis updates. This degradation of the analysis updates presumably explains the

occasional erroneous weakening of the model storm during the data assimilation process (e.g., after  $t = 30$ , 60, and 80 min in Fig. 15c). By alleviating the storm displacement errors, the FAT reduces the LETKF's sensitivity to these errors, potentially resulting in far superior analyses and forecasts of storms.

#### d. Observation-space diagnostics

To further evaluate the FAT's impact on analyses, observation-space diagnostic statistics, including mean innovation (MI), root-mean-square innovation (RMSI), and consistency ratio (CR), are computed from the

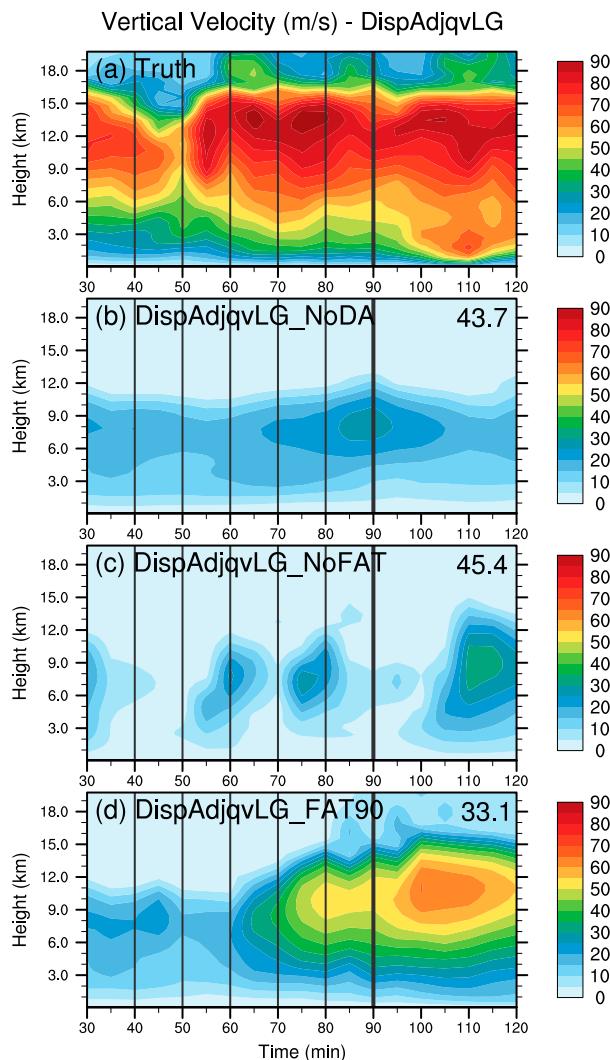


FIG. 15. As in Fig. 5, but for DispAdjvqLG.

reflectivity priors for the NoFAT and FAT90 ensembles in the DispTh, DispLG, and FastTh experiments. The priors in the FAT90 ensembles are taken from the phase-corrected reflectivity field. The statistics calculations are restricted to grid points where observed or forecast reflectivity is  $\geq 10$  dBZ. The equations for MI and CR can be found in Dowell and Wicker (2009), but it is important to note that the equation for RMSI, which is the denominator of the CR equation, is not that given in Dowell and Wicker [2009, Eq. (3.2)]—which is actually the innovation variance—but rather that given in Dowell et al. [2011; Eq. (4.1)].

Beginning with MI, applying the FAT prior to DA reduces the initial reflectivity bias by at least 5 dBZ in all three sets of experiments (Fig. 17a). At some data assimilation times, the NoFAT ensembles have a smaller bias than the FAT90 ensembles, but this lower bias

could be the result of LETKF expanding the area of reflectivity by introducing a new storm where the Truth storm exists while not removing the original displaced storm. Starting at  $t = 50$  min for DispLG,  $t = 60$  min for DispTh, and  $t = 70$  min for FastTh, the FAT90 ensembles exhibit a lower bias for the rest of the DA window, including a near-zero bias for DispTh\_FAT90 and FastTh\_FAT90 at  $t = 90$  min (Fig. 17a). Conversely, the NoFAT ensembles have biases  $>7$  dBZ prior to the final DA (Fig. 17a). The lower mean innovations for the FAT90 ensembles translate into substantially lower RMSI values, as compared to the NoFAT ensembles (i.e., 2–7 dBZ lower through the second half of the DA window; Fig. 17b).

Prior to the first DA cycle, the NoFAT ensembles have higher CR values (i.e.,  $\sim 0.5$ – $0.55$  vs  $\sim 0.3$ – $0.4$ ), owing to their having larger ensemble spread (not shown) than the FAT90 ensembles (Fig. 17c). This result is expected, since the first application of FAT reduces the ensemble spread. After the first or second data assimilations, however, the FAT90 ensembles begin to exhibit larger CR values than the NoFAT ensembles (Fig. 17c), owing to the smaller RMSI in the former. During the second half of the DA window, the CR values for DispTh\_FAT90 and FastTh\_FAT90 rapidly increase to values around 1, which indicates good ensemble dispersion (Fig. 17c). While CR values for DispLG\_FAT90 only increase to near 0.45 by the end of the DA window, they are still substantially larger than for DispLG\_NoFAT (i.e.,  $\sim 0.15$  difference; Fig. 17c). Overall, the FAT produces smaller innovations and therefore better ensemble dispersion. Qualitatively similar results were obtained for the other FAT variants and for the radial velocity priors as well (not shown).

### 5. Summary and discussion

Storm displacement errors can originate from biases in both parameterization schemes and the analyzed storm environment and are potentially exacerbated during data assimilation. The alleviation of these errors would directly benefit the WoF project’s goal of providing accurate storm-scale analyses and short-term forecasts of severe convection to meteorologists in an operational warning setting. As a potential step toward reducing storm displacement errors, we introduced a new version of the feature alignment technique (FAT) from Nehr Korn et al. (2014). In our version, the FAT minimizes a cost function between observed and forecast composite reflectivity fields to produce a 2D field of displacement-correcting vectors. Similar to Nehr Korn et al. (2014), our cost function is composed of a residual error term and four penalty terms, including smoothness, divergence, magnitude, and barrier.

Time: 95.0 min

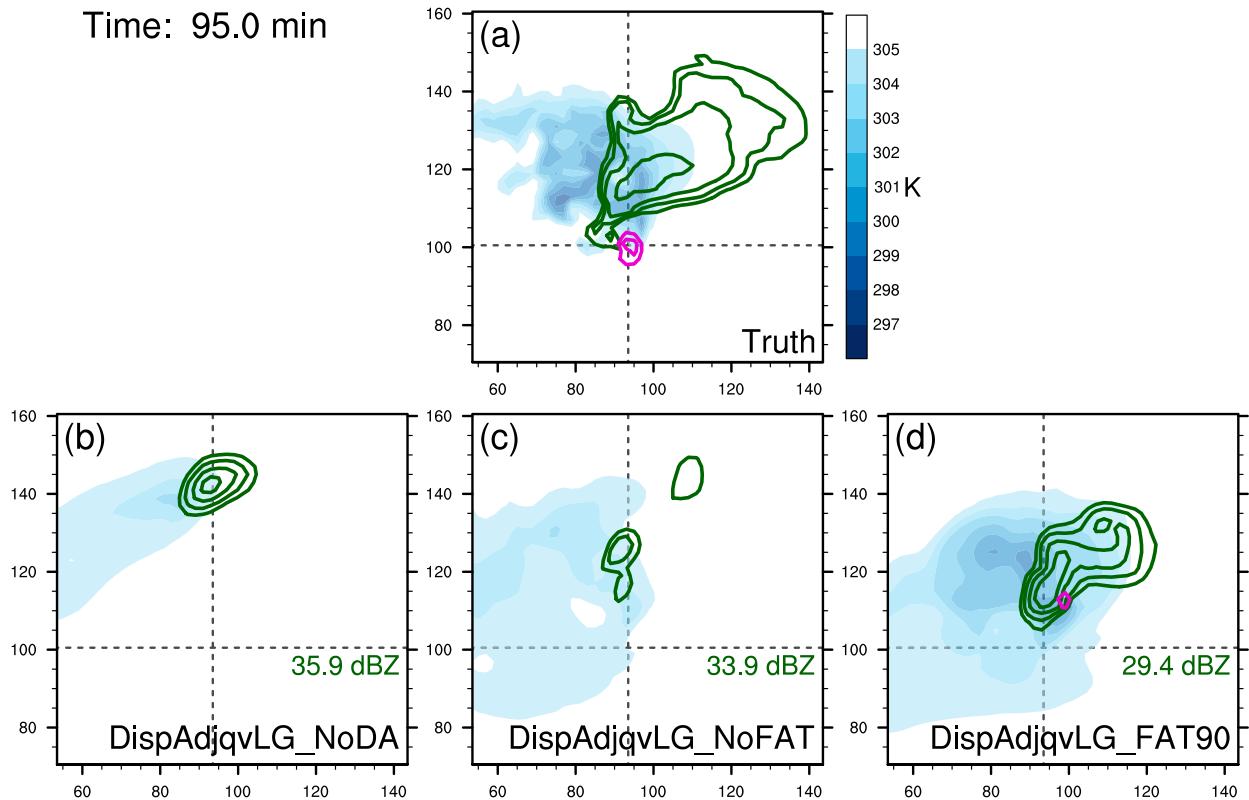


FIG. 16. As in Fig. 7, but for DispAdjqvLG.

In contrast to [Nehrkorn et al. \(2014\)](#) and other previous feature calibration and alignment (FCA) studies, who perform some or all of the minimization in spectral space, our minimization of the cost function is performed entirely in gridpoint space, so our version of the FAT incorporates additional smoothing techniques, including grid thinning and smoothing of the observed and forecast composite reflectivity fields.

The FAT was merged with the NSSL CM1–LETKF system as a standalone program that runs prior to each analysis update. A data assimilation OSSE framework was designed to assess the FAT’s impact on storm-scale analyses and short-term forecasts of an isolated supercell. A 2-h Truth simulation of an isolated supercell was generated on a 250-m grid and provided pseudo-radar observations and verification for the experiments. Four 2-km ensemble simulations without data assimilation (i.e., NoDA) and with mean storm initialization location, initialization sounding, and/or microphysics scheme differing from the Truth run were produced to provide backgrounds for the corresponding data assimilation experiments. Analyses and subsequent forecasts generated using either data assimilation alone or combined with one of four variations of the FAT were evaluated against the 250-m Truth run and 2-km NoDA ensembles.

In all of the OSSEs, the FAT successfully corrected location errors during the data assimilation window, thereby greatly improving storm intensity, spinup time, and structure without introducing detrimental model imbalances. These improvements persisted into the 30-min forecasts initialized from the final analysis. Forecast location and intensity results are summarized in [Fig. 18](#). These results reveal the substantial improvement in storm locations for all FAT ensembles while maintaining or improving storm intensity accuracy ([Fig. 18](#)). The LETKF-only (i.e., NoFAT) forecasts performed the worst for both metrics ([Fig. 18](#)). The FAT was particularly valuable in the dry boundary layer experiments, in which data assimilation alone produced only a weak storm. In experiments where an initial displacement was present but not reinforced throughout the data assimilation window, running the FAT only at the first analysis cycle sufficed to drastically improve analyses and forecasts. However, when a storm motion bias was introduced (via a stronger environmental wind profile) to increase displacement errors during each forecast cycle, running the FAT at every data assimilation step (rather than just the first one) produced more benefit.

In addition to the experiments shown in this study, we conducted experiments using the fast sounding with imperfect microphysics (i.e., FastLG) and experiments

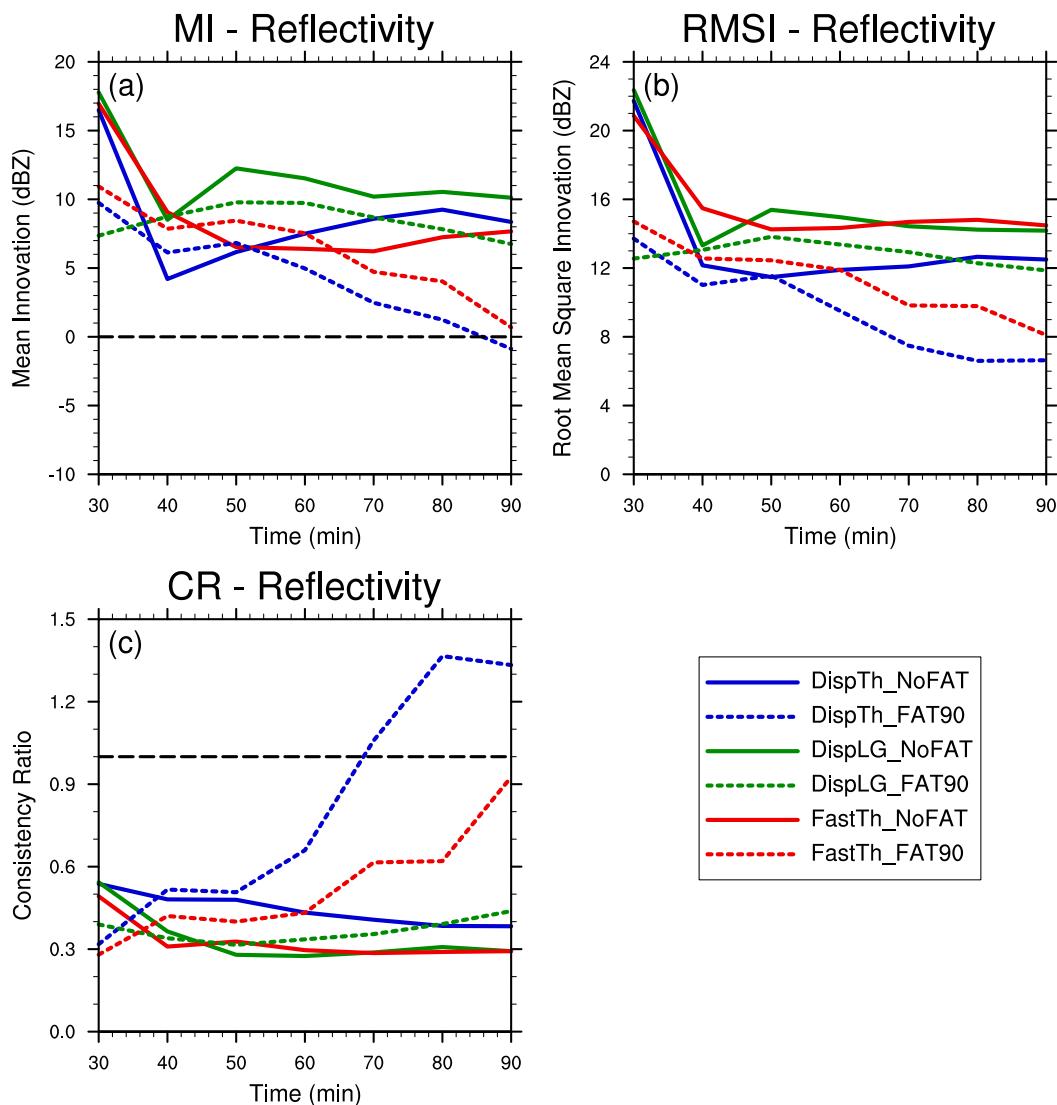


FIG. 17. Observation-space diagnostic statistics, including (a) MI (dBZ), (b) RMSI (dBZ), and (c) CR, are plotted for forecast reflectivity. The ensembles used include DispTh\_NoFAT (solid blue line), DispTh\_FAT90 (dashed blue line), DispLG\_NoFAT (solid green line), DispLG\_FAT90 (dashed green line), FastTh\_NoFAT (solid red line), and FastTh\_FAT90 (dashed red line).

combining the fast sounding with perfect and imperfect microphysics and with the Truth run's storm displaced northeastward to reduce the initial storm displacement from ~14 to ~8.5 km. These experiments yielded the same general results as those presented herein. Overall, the OSSE results suggest the FAT can improve analyses and short-term forecasts from real convection-allowing ensemble systems. More specifically, as shown by the experiments, correcting phase errors using the FAT during data assimilation can lead to a reduction in the northward and eastward bias commonly exhibited by convection-allowing forecast systems. Ideally, there would not be a need for storm-displacement correction techniques like the FAT.

However, storm-scale NWP phase errors are unlikely to dramatically improve in the foreseeable future, so the FAT or other similar error mitigation methods will continue to be valuable and needed for effective data assimilation.

In future studies, additional OSSEs will be conducted using events with multiple supercells, mesoscale convective systems, and mixed modes. Since forecasts in the experiments herein were only produced out to 30 min, future experiments will explore how long the FAT's beneficial impact in forecasts lasts. Beyond that, the FAT will be tested in real-data frameworks with full model physics using isolated supercell cases to start and eventually more complex (e.g., mixed mode) events. These

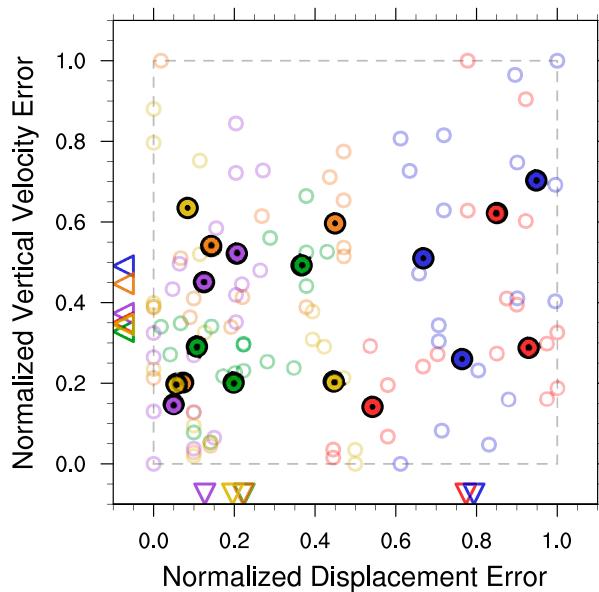


FIG. 18. Summary plot comparing displacement errors and vertical velocity errors between the Truth run and ensembles during the free forecast. The errors are normalized using maximum and minimum values from all forecasts (i.e., six output times) for each experiment set (e.g., DispTh), so a value of 0 (1) is the best (worst). Hollowed circles indicate the NoDA (red), NoFAT (blue), FAT90 (green), FAT50 (purple), FAT3gp (orange), and FAT1st (gold) ensemble forecast errors for the DispTh, DispLG, and FastTh experiment sets. Filled circles indicate the average of these errors within each experiment set. The average errors across the three experiment sets are represented by the triangles on the bottom (displacement error) and left (vertical velocity error) sides of the plot. The gray dashed lines represent the upper and lower limits of the normalized errors.

real-data experiments will be performed using a forecast system similar to the NSSL Experimental WoF System for ensembles (NEWS-e; Wheatley et al. 2015; Jones et al. 2016). Before those real-data experiments can be completed, terrain effects will need to be incorporated into the FAT as in Nehr Korn et al. (2014). Throughout these future tests, the tunable FAT parameters will need to be tuned to static or dynamic values based on sensitivities to storm mode and geographic location. The FAT could ultimately provide great value to an operational Warn-on-Forecast ensemble system.

*Acknowledgments.* This research was performed while the author held an NRC Research Associateship award at the National Severe Storms Laboratory. Funding for the second author was provided by NOAA/Office of Oceanic and Atmospheric Research under NOAA–University of Oklahoma Cooperative Agreement NA11OAR4320072, U.S. Department of Commerce. We also thank Patrick Skinner for helpful discussions and for informally reviewing an early version of the manuscript. Valuable local computing assistance was provided by Gerry Creager,

Jesse Butler, and Jeff Horn. The authors thank the three anonymous reviewers for their comments and suggestions, which improved the manuscript.

#### REFERENCES

- Anderson, J., T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Avellano, 2009: The Data Assimilation Research Testbed: A community facility. *Bull. Amer. Meteor. Soc.*, **90**, 1283–1296, <https://doi.org/10.1175/2009BAMS2618.1>.
- Barker, D., and Coauthors, 2012: The Weather Research and Forecasting Model's Community Variational/Ensemble Data Assimilation System: WRFDA. *Bull. Amer. Meteor. Soc.*, **93**, 831–843, <https://doi.org/10.1175/BAMS-D-11-00167.1>.
- Braun, S. A., and W. Tao, 2000: Sensitivity of high-resolution simulations of Hurricane Bob (1991) to planetary boundary layer parameterizations. *Mon. Wea. Rev.*, **128**, 3941–3961, [https://doi.org/10.1175/1520-0493\(2000\)129<3941:SOHRSO>2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)129<3941:SOHRSO>2.0.CO;2).
- Brewster, K. A., 2003a: Phase-correcting data assimilation and application to storm-scale numerical weather prediction. Part I: Method description and simulation testing. *Mon. Wea. Rev.*, **131**, 480–492, [https://doi.org/10.1175/1520-0493\(2003\)131<0480:PCDAAA>2.0.CO;2](https://doi.org/10.1175/1520-0493(2003)131<0480:PCDAAA>2.0.CO;2).
- , 2003b: Phase-correcting data assimilation and application to storm-scale numerical weather prediction. Part II: Application to a severe storm outbreak. *Mon. Wea. Rev.*, **131**, 493–507, [https://doi.org/10.1175/1520-0493\(2003\)131<0493:PCDAAA>2.0.CO;2](https://doi.org/10.1175/1520-0493(2003)131<0493:PCDAAA>2.0.CO;2).
- Bryan, G. H., and J. M. Fritsch, 2002: A benchmark simulation for moist nonhydrostatic numerical models. *Mon. Wea. Rev.*, **130**, 2917–2928, [https://doi.org/10.1175/1520-0493\(2002\)130<2917:ABSFMN>2.0.CO;2](https://doi.org/10.1175/1520-0493(2002)130<2917:ABSFMN>2.0.CO;2).
- Coniglio, M. C., 2012: Verification of RUC 0–1-h forecasts and SPC mesoscale analyses using VORTEX2 soundings. *Wea. Forecasting*, **27**, 667–683, <https://doi.org/10.1175/WAF-D-11-00096.1>.
- , J. Correia, P. T. Marsh, and F. Kong, 2013: Verification of convection-allowing WRF Model forecasts of the planetary boundary layer using sounding observations. *Wea. Forecasting*, **28**, 842–862, <https://doi.org/10.1175/WAF-D-12-00103.1>.
- Dawson, D. T., II, M. Xue, J. A. Milbrandt, and A. Shapiro, 2015: Sensitivity of real-data simulations of the 3 May 1999 Oklahoma City tornadic supercell and associated tornadoes to multimoment microphysics. Part I: Storm- and tornado-scale numerical forecasts. *Mon. Wea. Rev.*, **143**, 2241–2265, <https://doi.org/10.1175/MWR-D-14-00279.1>.
- Deardorff, J. W., 1980: Stratocumulus-capped mixed layers derived from a three-dimensional model. *Bound.-Layer Meteor.*, **18**, 495–527, <https://doi.org/10.1007/BF00119502>.
- Dowell, D. C., and L. J. Wicker, 2009: Additive noise for storm-scale ensemble data assimilation. *J. Atmos. Oceanic Technol.*, **26**, 911–927, <https://doi.org/10.1175/2008JTECHA1156.1>.
- , —, and C. Snyder, 2011: Ensemble Kalman filter assimilation of radar observations of the 8 May 2003 Oklahoma City supercell: Influences of reflectivity observations on storm-scale analyses. *Mon. Wea. Rev.*, **139**, 272–294, <https://doi.org/10.1175/2010MWR3438.1>.
- Ebert, E. E., 2001: Ability of a poor man's ensemble to predict the probability and distribution of precipitation. *Mon. Wea. Rev.*, **129**, 2461–2480, [https://doi.org/10.1175/1520-0493\(2001\)129<2461:AOAPMS>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<2461:AOAPMS>2.0.CO;2).
- Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10 143–10 162, <https://doi.org/10.1029/94JC00572>.

- Grassotti, C., H. Iskenderian, and R. N. Hoffman, 1999: Calibration and alignment. *J. Appl. Meteor.*, **38**, 677–695, [https://doi.org/10.1175/1520-0450\(1999\)038<0677:CAA>2.0.CO;2](https://doi.org/10.1175/1520-0450(1999)038<0677:CAA>2.0.CO;2).
- Hoffman, R. N., and C. Grassotti, 1996: A technique for assimilating SSM/I observations of marine atmospheric storms: Tests with ECMWF analyses. *J. Appl. Meteor.*, **35**, 1177–1188, [https://doi.org/10.1175/1520-0450\(1996\)035<1177:ATFASO>2.0.CO;2](https://doi.org/10.1175/1520-0450(1996)035<1177:ATFASO>2.0.CO;2).
- , Z. Liu, J.-F. Louis, and C. Grassotti, 1995: Distortion representation of forecast errors. *Mon. Wea. Rev.*, **123**, 2758–2770, [https://doi.org/10.1175/1520-0493\(1995\)123<2758:DROFE>2.0.CO;2](https://doi.org/10.1175/1520-0493(1995)123<2758:DROFE>2.0.CO;2).
- Hsiao, L.-F., C.-S. Liou, T.-C. Yeh, Y.-R. Guo, D.-S. Chen, K.-N. Huang, C.-T. Terng, and J.-H. Chen, 2010: A vortex relocation scheme for tropical cyclone initialization in Advanced Research WRF. *Mon. Wea. Rev.*, **138**, 3298–3315, <https://doi.org/10.1175/2010MWR3275.1>.
- Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D*, **230**, 112–126, <https://doi.org/10.1016/j.physd.2006.11.008>.
- Jones, T. A., D. Stensrud, L. Wicker, P. Minnis, and R. Palikonda, 2015: Simultaneous radar and satellite data storm-scale assimilation using an ensemble Kalman filter approach for 24 May 2011. *Mon. Wea. Rev.*, **143**, 165–194, <https://doi.org/10.1175/MWR-D-14-00180.1>.
- , K. Knopfmeier, D. Wheatley, G. Creager, P. Minnis, and R. Palikonda, 2016: Storm-scale data assimilation and ensemble forecasting with the NSSL Experimental Warn-on-Forecast System. Part II: Combined radar and satellite data experiments. *Wea. Forecasting*, **31**, 297–327, <https://doi.org/10.1175/WAF-D-15-0107.1>.
- Kain, J. S., and Coauthors, 2008: Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP. *Wea. Forecasting*, **23**, 931–952, <https://doi.org/10.1175/WAF2007106.1>.
- Knievel, J. C., G. H. Bryan, and J. P. Hacker, 2007: Explicit numerical diffusion in the WRF Model. *Mon. Wea. Rev.*, **135**, 3808–3824, <https://doi.org/10.1175/2007MWR2100.1>.
- Lehmiller, G. S., H. B. Bluestein, P. J. Neiman, F. M. Ralph, and W. F. Feltz, 2001: Wind structure in a supercell thunderstorm as measured by a UHF wind profiler. *Mon. Wea. Rev.*, **129**, 1968–1986, [https://doi.org/10.1175/1520-0493\(2001\)129<1968:WSIAST>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<1968:WSIAST>2.0.CO;2).
- Miyoshi, T., 2011: The Gaussian approach to adaptive covariance inflation and its implementation with the local ensemble transform Kalman filter. *Mon. Wea. Rev.*, **139**, 1519–1535, <https://doi.org/10.1175/2010MWR3570.1>.
- Nehrkorn, T., B. Woods, T. Auligné, and R. N. Hoffman, 2014: Application of feature calibration and alignment to high-resolution analysis: Examples using observations sensitive to cloud and water vapor. *Mon. Wea. Rev.*, **142**, 686–702, <https://doi.org/10.1175/MWR-D-13-00164.1>.
- , —, R. N. Hoffman, and T. Auligné, 2015: Correcting for position errors in variational data assimilation. *Mon. Wea. Rev.*, **143**, 1368–1381, <https://doi.org/10.1175/MWR-D-14-00127.1>.
- Polak, E., and G. Ribière, 1969: Note sur la convergence de méthodes de directions conjuguées. *ESAIM: Mathematical Modelling and Numerical Analysis—Modélisation Mathématique et Analyse Numérique*, **3**, 35–43, <http://eudml.org/doc/193115>.
- Potvin, C. K., and M. L. Flora, 2015: Sensitivity of idealized supercell simulations to horizontal grid spacing: Implications for Warn-on-Forecast. *Mon. Wea. Rev.*, **143**, 2998–3024, <https://doi.org/10.1175/MWR-D-14-00416.1>.
- Shen, Y., and G. Zha, 2010: Improvement of the WENO scheme smoothness estimator. *Int. J. Numer. Methods Fluids*, **64**, 653–675, <https://doi.org/10.1002/fld.2168>.
- Shu, C.-W., 2003: High-order finite difference and finite volume WENO schemes and discontinuous Galerkin methods for CFD. *Int. J. Comput. Fluid Dyn.*, **17**, 107–118, <https://doi.org/10.1080/1061856031000104851>.
- Snyder, C., and F. Zhang, 2003: Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.*, **131**, 1663–1677, <https://doi.org/10.1175/2555.1>.
- Stensrud, D. J., and Coauthors, 2009: Convective-scale Warn-on-Forecast System. *Bull. Amer. Meteor. Soc.*, **90**, 1487–1500, <https://doi.org/10.1175/2009BAMS2795.1>.
- , and Coauthors, 2013: Progress and challenges with Warn-on-Forecast. *Atmos. Res.*, **123**, 2–16, <https://doi.org/10.1016/j.atmosres.2012.04.004>.
- Stratman, D. R., and K. A. Brewster, 2017: Sensitivities of 1-km forecasts of 24 May 2011 tornadic supercells to microphysics parameterizations. *Mon. Wea. Rev.*, **145**, 2697–2721, <https://doi.org/10.1175/MWR-D-16-0282.1>.
- Supinie, T. A., N. Yussouf, Y. Jung, M. Xue, J. Cheng, and S. Wang, 2017: Comparison of the analyses and forecasts of a tornadic supercell storm from assimilating phased-array radar and WSR-88D observations. *Wea. Forecasting*, **32**, 1379–1401, <https://doi.org/10.1175/WAF-D-16-0159.1>.
- Tanamachi, R. L., P. L. Heinselman, and L. J. Wicker, 2015: Impacts of a storm merger on the 24 May 2011 El Reno, Oklahoma, tornadic supercell. *Wea. Forecasting*, **30**, 501–524, <https://doi.org/10.1175/WAF-D-14-00164.1>.
- Thompson, G., P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Mon. Wea. Rev.*, **136**, 5095–5115, <https://doi.org/10.1175/2008MWR2387.1>.
- Thompson, T. E., L. J. Wicker, X. Wang, and C. Potvin, 2015: A comparison between the local ensemble transform Kalman filter and the ensemble square root filter for the assimilation of radar data in convective-scale models. *Quart. J. Roy. Meteor. Soc.*, **141**, 1163–1176, <https://doi.org/10.1002/qj.2423>.
- Wheatley, D. M., N. Yussouf, and D. J. Stensrud, 2014: Ensemble Kalman filter analyses and forecasts of a severe mesoscale convective system using different choices of microphysics schemes. *Mon. Wea. Rev.*, **142**, 3243–3263, <https://doi.org/10.1175/MWR-D-13-00260.1>.
- , K. H. Knopfmeier, T. A. Jones, and G. J. Creager, 2015: Storm-scale data assimilation and ensemble forecasting with the NSSL Experimental Warn-on-Forecast System. Part I: Radar data experiments. *Wea. Forecasting*, **30**, 1795–1817, <https://doi.org/10.1175/WAF-D-15-0043.1>.
- Xue, M., Y. Jung, and G. Zhang, 2007: Error modeling of simulated reflectivity observations for ensemble Kalman filter assimilation of convective storms. *Geophys. Res. Lett.*, **34**, L10802, <https://doi.org/10.1029/2007GL029945>.
- , M. Hu, and A. D. Schenkman, 2014: Numerical prediction of the 8 May 2003 Oklahoma City tornadic supercell and embedded tornado using ARPS with the assimilation of WSR-88D data. *Wea. Forecasting*, **29**, 39–62, <https://doi.org/10.1175/WAF-D-13-00029.1>.
- Yussouf, N., E. R. Mansell, L. J. Wicker, D. M. Wheatley, and D. J. Stensrud, 2013: The ensemble Kalman filter analyses and forecasts of the 8 May 2003 Oklahoma City tornadic supercell storm using single- and double-moment microphysics schemes. *Mon. Wea. Rev.*, **141**, 3388–3412, <https://doi.org/10.1175/MWR-D-12-00237.1>.
- , D. C. Dowell, L. J. Wicker, K. H. Knopfmeier, and D. M. Wheatley, 2015: Storm-scale data assimilation and ensemble forecasts for the 27 April 2011 severe weather outbreak in Alabama. *Mon. Wea. Rev.*, **143**, 3044–3066, <https://doi.org/10.1175/MWR-D-14-00268.1>.